TRIAL STATISTICAL ANALYSIS PLAN

c37199574-02

BI Trial No.:	1403-0008
Title:	Brightline-1: A Phase II/III, randomized, open-label, multi-center study of brigimadlin (BI 907828) compared to doxorubicin as first line treatment of patients with advanced dedifferentiated liposarcoma
Investigational Product(s):	Brigimadlin (BI 907828)
Responsible trial statistician(s):	[redacted] Phone: [redacted]
Original date of statistical analysis plan:	01 FEB 2023
Current version and date:	Version 2.0, 19 APR 2024

1. TABLE OF CONTENTS

LIST OF TABLES

2. LIST OF ABBREVIATIONS

Term	Definition / description
ALT	Alanine Aminotransferase
ANC	Absolute Neutrophil Counts
AQA	Analgesic Quantification Algorithm
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical
AUC	Area Under the Curve
AE	Adverse Event
BI	Boehringer Ingelheim
BIRDS	BI Regulatory Documents for Submission
BMI	Body Mass Index
BRAVE	BI Rave
BRPM	Blinded Report Planning Meeting
BSA	Body Surface Area
CI	Confidence Interval
cm	Centimeters
C_{max}	Maximum Plasma Concentration
CR	Complete Response
CRF	Case Report Form
CS	Clinical Significance
CTC	Common Terminology Criteria
CTP	Clinical Trial Protocol
CTR	Clinical Trial Report
DC	Disease Control
DDLPS	Dedifferentiated Liposarcoma
DDR	Data Delivery Request
DMC	Data Monitoring Committee
DOR	Duration of Objective Response
ECG	Electrocardiogram
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
EDMS	Electronic Document Management System

Term	Definition / description
EORTC	European Organization for Research and Treatment on Cancer
EOT	End of Treatment
FA	Fatigue
FACT-GP5	Functional Assessment of Cancer Therapy – Global Physical item #5
GCSF	Granulocyte Colony-Stimulating Factor
GRL	Global Regulatory Lead
gRS	Global Randomization Specialist
HR	Hazard Ratio
HRQoL	Health Related Quality of Life
HSAP	Health Economics Statistical Analysis Plan
ICH	International Conference on Harmonisation
iPD	Important Protocol Deviation
IRT	Interactive Response Technology
iSTAT	Independent Statistician
ITT	Intention-to-Treat
kg	Kilograms
L	Liter
m	Meters
mm	Millimeters
Max	Maximum
MCID	Minimal Clinically Important Difference
MDM2	Mouse Double Minute 2 Homolog
MedDRA	Medical Dictionary for Regulatory Activities
Min	Minimum
mmol	Millimole
NCA	Non-Compartmental Analysis
NE	Not Evaluable
OR	Objective Response
ORR	Objective Response Rate
OS	Overall Survival
PA	Pain
PD	Progressive Disease

Term	Definition / description
PF	Physical Functioning
PFS	Progression-Free Survival
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
PK	Pharmacokinetics
PK/PD	Pharmacokinetics/Pharmacodynamics
PR	Partial Response
PRO	Patient Reported Outcome
PRO-CTCAE	PRO version of the Common Terminology Criteria for Adverse Events
PROM	PRO Measure
PT	Preferred Term
QL	Quality of Life
QLQ-C30	Quality of Life Core Questionnaire 30
RCI	Repeated Confidence Interval
RDI	Relative Dose Intensity
RECIST	Response Evaluation Criteria in Solid Tumours
REP	Residual Effect Period
RPSFTM	Rank-Preserving Structural Failure Time Model
RS	Raw Score
RUN	Data Ready to be Unblinded and/or Final Trial Closure Notification
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Stable Disease
SIR	Synoptic Interim Report
SMQ	Standardized MedDRA Query
SOC	System Organ Class
TIR	Treatment Information Release
TMCP	Translational Medicine and Clinical Pharmacology
TMF	Trial Master File
TSAP	Trial Statistical Analysis Plan
UDAEC	User-Defined Adverse Event Category
ULN	Upper Limit of Normal

Term	Definition / description
VAS	Visual Analogue Scale
WBC	White Blood Cell
WDLPS	Well-Differentiated Liposarcoma
WHO DD	World Health Organization Drug Dictionary

3. INTRODUCTION

As per International Conference on Harmonisation (ICH) E9 guidance (1), the purpose of this document is to provide a more technical and detailed elaboration of the principal features of the analysis described in the protocol and to include detailed procedures for executing the statistical analysis of the primary and secondary variables and other data.

This Trial Statistical Analysis Plan (TSAP) assumes familiarity with the Clinical Trial Protocol (CTP), including protocol amendments. In particular, the TSAP is based on the planned analysis specification as written in CTP Section 7 “Statistical Methods and Determination of Sample Size”. Therefore, TSAP readers may consult the CTP for more background information on the study, e.g., trial objectives, trial design and population, treatments, definition of measurements and variables, planning of sample size, randomization, etc.

This document contains details for the following analyses:

- The interim futility analysis at approximately the same time as the end of Phase II;
- The interim Clinical Trial Report (CTR) for the primary analysis of the primary endpoint progression-free survival (PFS) (based on blinded central independent review) in Phase III;
- The final CTR for the primary analysis of the selected secondary endpoint overall survival (OS) at the end of Phase III (i.e., at the end of the trial).

SAS® Version 9.4 (or later) will be used for all analyses.

Details for the interim analysis for dose selection in the Phase II part of the trial are described in CTP Section 7.2.8 as well as in the Data Monitoring Committee (DMC) SAP and a separate PK/PD SAP.

Analyses of the biomarker data are described in a separate biomarker SAP.

4. CHANGES IN THE PLANNED ANALYSIS OF THE STUDY

Due to the trial enrolling patients much faster than anticipated, the trial design, number of patients, and statistical methods have been updated to reflect a larger overall sample size and PFS event rate. These updates are reflected in the changes to the CTP as of CTP version 4.0.

5. ENDPOINTS

All endpoints are stated and comprehensively defined in Section 2 of the CTP.

5.1 PRIMARY ENDPOINT

The primary endpoint of the trial is progression-free survival (PFS) based on blinded central independent review. PFS is defined as the time interval from randomization until the earliest of tumor progression according to RECIST version 1.1 (based on blinded central independent review) or death from any cause, whichever occurs first.

The primary endpoint of PFS (based on blinded central independent review) will be assessed and compared between the selected investigational arm and the doxorubicin control arm at an interim futility analysis at approximately the same time as the end of Phase II and at a primary analysis of PFS during Phase III.

Derivations below are described in days. However, the endpoints below will be presented in months in the statistical tables produced for the interim futility analysis and the CTR. Formulas for converting days to months are in Section 7.

For patients with known date of progression (based on blinded central independent review) or death:

- PFS [days] = earlier date of progression (based on blinded central independent review) or death – date of randomization + 1.

For patients who will be censored:

- PFS (censored) [days] = date of last imaging assessment without disease progression (based on blinded central independent review) – date of randomization + 1.

Rules to determine events and censoring for the interim futility analysis of PFS at approximately the same time as the end of Phase II, and for the primary analysis of PFS during Phase III, are in Table 7.3: 1 of the CTP.

As described in the CTP, patients administered a subsequent anti-cancer therapy recorded in the case report form (CRF) but not presenting with progressive disease (PD) before the start of the new anti-cancer therapy will be censored at the date of last imaging before the new anti-cancer therapy started. For patients in the doxorubicin arm who cross-over to brigimadlin, start of brigimadlin is considered start of subsequent anti-cancer therapy, and the corresponding censoring rules will be applied if applicable.

All randomized patients in the selected investigational arm and the doxorubicin control arm will be included in the interim futility analysis of PFS at approximately the same time as the end of Phase II and in the primary analysis of PFS during Phase III.

Patients randomized to the unselected dose of brigimadlin who later change to the selected dose of brigimadlin after dose selection will be censored at the date of the patient's last tumor assessment before the dose change occurs if there is no PFS event prior to the dose change.

If a patient has a radiological examination performed over a number of days, i.e., target lesions assessed on day x and non-target lesions assessed on day y, a single date for the derivation of PFS for each patient is still required. The following rules will be applied:

- If the overall response is PD, the earliest date of multiple assessments will be taken.
- If the overall response is SD, PR, or CR, the latest of multiple assessment dates will be taken, i.e., in the case of SD, PR, or CR, the latest of multiple dates will be used to censor the patient.

If a patient does not have an image in one of the time windows described in Table 6.7: 1, then the patient will be considered to have "missed an assessment" for that time point.

An image where an overall response of NE (not evaluable) has been assigned is not considered evaluable. Images for which NE is assigned as the overall response at an imaging time point are considered to be missed assessments. It cannot be assumed that NE indicates progression has not occurred.

5.1.1 Derivations of central independent review data with adjudication

Central independent review of images is performed using a "two-plus-one" configuration, i.e., for each patient, two radiologists (i.e., Readers 1 and 2) separately assess the patient's images from screening (i.e., baseline) and all subsequent imaging visits to determine time point responses according to RECIST version 1.1. A third reviewer (i.e., Adjudicator) selects the preferred set of assessments only when Readers 1 and 2 have discordant Overall Response assessments for a given time point.

More detailed attributes of the adjudication process in the trial are as follows:

- When Readers 1 and 2 disagree as to whether an Overall Response of PD has occurred for a given time point, the case is immediately assigned for adjudication. Per protocol, an individual patient must discontinue the current treatment (of brigimadlin or doxorubicin) upon confirmed PD by central independent review.
- When Readers 1 and 2 disagree for an Overall Response not related to PD (i.e., any other discrepancy in the Overall Response) for a given time point, the case is not immediately assigned for adjudication. Adjudication will be performed: when the patient discontinues from the trial, at the end of the trial, or prior to data transfers for interim data analysis upon agreement by the sponsor and central imaging vendor.
- For all adjudications, the Adjudicator selects one of the two complete sets of assessments, whichever one the Adjudicator considers to be a more accurate representation of the patient's burden of disease and response to treatment.

- It is possible that adjudication may occur more than once per patient.
- It is possible that adjudication never occurs for a patient. This scenario occurs when Readers 1 and 2 have concordant Overall Responses at every time point.

Additional details about the central independent review process in the trial are in the Independent Review Charter. Considering the adjudication rules in the trial, endpoints based on blinded central independent review data are derived as described in the Analysis Dataset Specification plan.

5.2 SECONDARY ENDPOINTS

5.2.1 Key secondary endpoint

No key secondary endpoints are defined for the trial.

5.2.2 Secondary endpoints

Objective response (OR) based on *blinded central independent review*

OR is defined as a best overall response of confirmed complete response (CR) or confirmed partial response (PR) according to RECIST version 1.1 based on *blinded central independent review* from the date of randomization until disease progression (based on blinded central independent review), death, last evaluable tumor assessment before start of subsequent anti-cancer therapy, loss to follow-up, or withdrawal of consent, whichever occurs first. Each patient will be assigned as having a best overall response of one of the following RECIST categories: CR, PR, SD (stable disease), PD, or NE.

A best overall response of CR or PR must be confirmed by a subsequent tumor assessment (based on blinded central independent review), i.e., a confirmed response requires a repeat observation at least 4 weeks apart.

Unconfirmed PR or CR will be presented as a sub-category of SD. SD must last for at least 42 days to be presented as SD. SD lasting less than 42 days and followed by PD will be presented as a sub-category of PD. SD lasting less than 42 days and without evaluable response thereafter will be listed as “Not evaluable”. Any assessments done in compliance with the pre-specified (-)7-day assessment window in the CTP will also be accounted for as within the 42-day duration.

For patients who die, best overall response will be calculated based on data up until the last evaluable RECIST assessment prior to death. Death will not contribute as PD for best overall response. Patients who die prior to an evaluable imaging assessment will contribute as NE.

To be aligned with the primary endpoint derivation, tumor assessments after two or more consecutively missed assessments will not be considered when determining best overall response. In addition, to be aligned with the primary endpoint derivation, for patients randomized to the unselected dose of brigimadlin who later change to the selected dose of brigimadlin after dose selection, tumor assessments after the dose change occurs will not be considered when determining best overall response.

Rules for determining confirmed best overall response are provided in Table 5.2.2: 1.

Table 5.2.2: 1 Best overall response confirmation rules

Overall response (first time point)	Overall response (subsequent time point)	Confirmed best overall response
CR	CR	CR
CR	PD	SD as long as ≥ 42 days[b], otherwise PD
CR	NE/Missing[a]	SD as long as ≥ 42 days[b], otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD as long as ≥ 42 days[b], otherwise PD
PR	NE/Missing[a]	SD as long as ≥ 42 days[b], otherwise NE
SD	NE/Missing	SD as long as ≥ 42 days[b], otherwise NE
NE	NE/Missing	NE

a) If a patient has an overall response sequence of CR, NE/Missing, CR, then the patient's confirmed best overall response is CR. If a patient has an overall response sequence of PR, NE/Missing, PR, then the patient's confirmed best overall response is PR.

b) Any assessments done in compliance with the pre-specified (-)7-day assessment window in the CTP will also be accounted for as within the 42-day duration.

Overall survival (OS)

OS is defined as the time interval from randomization until death from any cause.

For patients with known date of death of any cause:

- OS [days] = date of death – date of randomization + 1.

For patients who will be censored:

- OS (censored) [days] = date of last contact when the patient is known to be alive – date of randomization + 1.

Date when the patient is last known to be alive will be derived using available data sources.

Rules to determine events and censoring for OS are provided in Table 5.2.2: 2.

Partial death dates will be imputed as described in Section 6.6.

Table 5.2.2: 2 Censoring rules for OS

Status at time of analysis	Outcome (Event or Censored)	Date of outcome
Patient died, and the date of death is known	Event	Date of death
Patient died, and the date of death is unknown	Event	Date of last contact when the patient is known to be alive + 1 day
Patient alive	Censored	Date of last contact when the patient is known to be alive
Unknown	Censored	Date of last contact when the patient is known to be alive

Duration of objective response (DOR) based on *blinded central independent review*

DOR is defined as the time interval from first documented confirmed OR based on *blinded central independent review* until disease progression (based on blinded central independent review) or death among patients with confirmed OR based on blinded central independent review, whichever occurs first.

DOR can only be calculated for patients with confirmed OR (based on blinded central independent review).

For patients with confirmed OR and disease progression (based on blinded central independent review) or death:

- DOR [days] = date of radiological assessment of progression (based on blinded central independent review) or date of death – date of first assessment indicating OR (based on blinded central independent review) + 1.

For patients with confirmed OR and without disease progression (based on blinded central independent review) or death:

- DOR (censored) [days] = date of last radiological assessment – date of first assessment indicating OR (based on blinded central independent review) + 1.

Rules to determine events and censoring for DOR are provided in Table 5.2.2: 3.

Only radiological assessments after first assessment indicating OR (based on blinded central independent review) should be taken into consideration. For patients who are randomized to the unselected dose level and change to the selected dose (after dose selection), only radiological assessments before the dose change will be taken into consideration.

Table 5.2.2: 3 Censoring rules for DOR

Situation	Outcome (Event or Censored)	Date of outcome
No other anti-cancer therapy		
Alive and not progressed, and no more than one consecutively missed radiological assessment	Censored	Date of last radiological assessment
Alive and not progressed, and two or more consecutively missed radiological assessments	Censored	Date of last radiological assessment prior to missed radiological assessments
Progressed, and zero or one missed radiological assessment prior to progression	Event	Date of radiological assessment of progression
Progressed, and two or more consecutively missed radiological assessments prior to progression	Censored	Date of last radiological assessment prior to missed assessments
Death but no progression, and zero or one missed radiological assessment prior to death	Event	Date of death
Death but no progression, and two or more consecutively missed radiological assessments prior to death	Censored	Date of last radiological assessment prior to missed assessments
Initiation of subsequent anti-cancer therapy		
Subsequent anti-cancer therapy started before progression or death, and no more than one consecutively missed radiological assessments prior to start of subsequent anti-cancer therapy	Censored	Date of last radiological assessment before initiation of subsequent anti-cancer therapy
Subsequent anti-cancer therapy started before progression or death, and two or more consecutively missed radiological assessments prior to start of subsequent anti-cancer therapy	Censored	Date of last radiological assessment prior to missed assessments

Disease control (DC) based on *blinded central independent review*

DC is defined as a best overall response (as detailed for OR based on blinded central independent review) of CR, PR, or SD according to RECIST version 1.1 based on *blinded central independent review*. A patient must have SD (based on blinded central independent

review) for at least 42 days to be counted as having DC. Any assessments done in compliance with the pre-specified (-)7-day assessment window in the CTP will also be accounted for as within the 42-day duration.

Health-Related Quality of Life (HRQoL)

HRQoL is assessed based on data collected through specific questionnaires, i.e., Patient Reported Outcome Measures (PROMs), from baseline until disease progression (based on blinded central independent review), death, start of subsequent therapy, lost to follow-up, or withdrawal of consent, whichever occurs first. The main PRO concepts of interest that will be considered for secondary PRO endpoints are physical functioning, the key symptoms related to liposarcoma and their impact (fatigue, fatigability, pain, fatigue impact, and pain impact), and overall HRQoL. The comparison between the investigational arm and the doxorubicin arm will be primarily evaluated at Week 18, as this is when most participants of the doxorubicin arm are expected to have reached the maximum cumulative dose of 450 mg/m^2. Another evaluation will be performed at Week 6 to capture early differences between treatment arms.

The PRO concepts will be assessed using the scores calculated from the data collected through the following questionnaires as described in Section 5.6.1 of the CTP at the time points specified in the Flowcharts in the CTP.

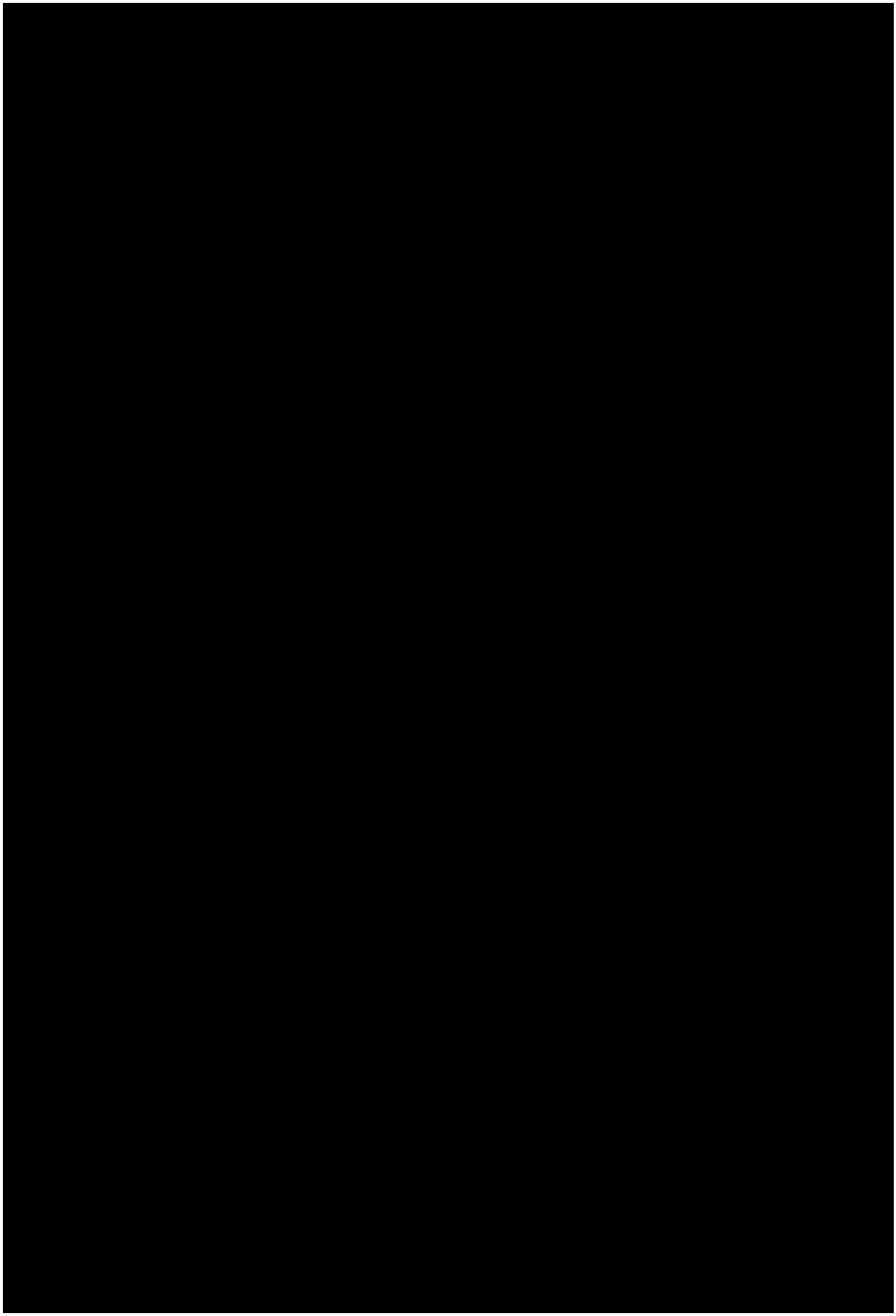
Hence, the HRQoL PRO measurements are:

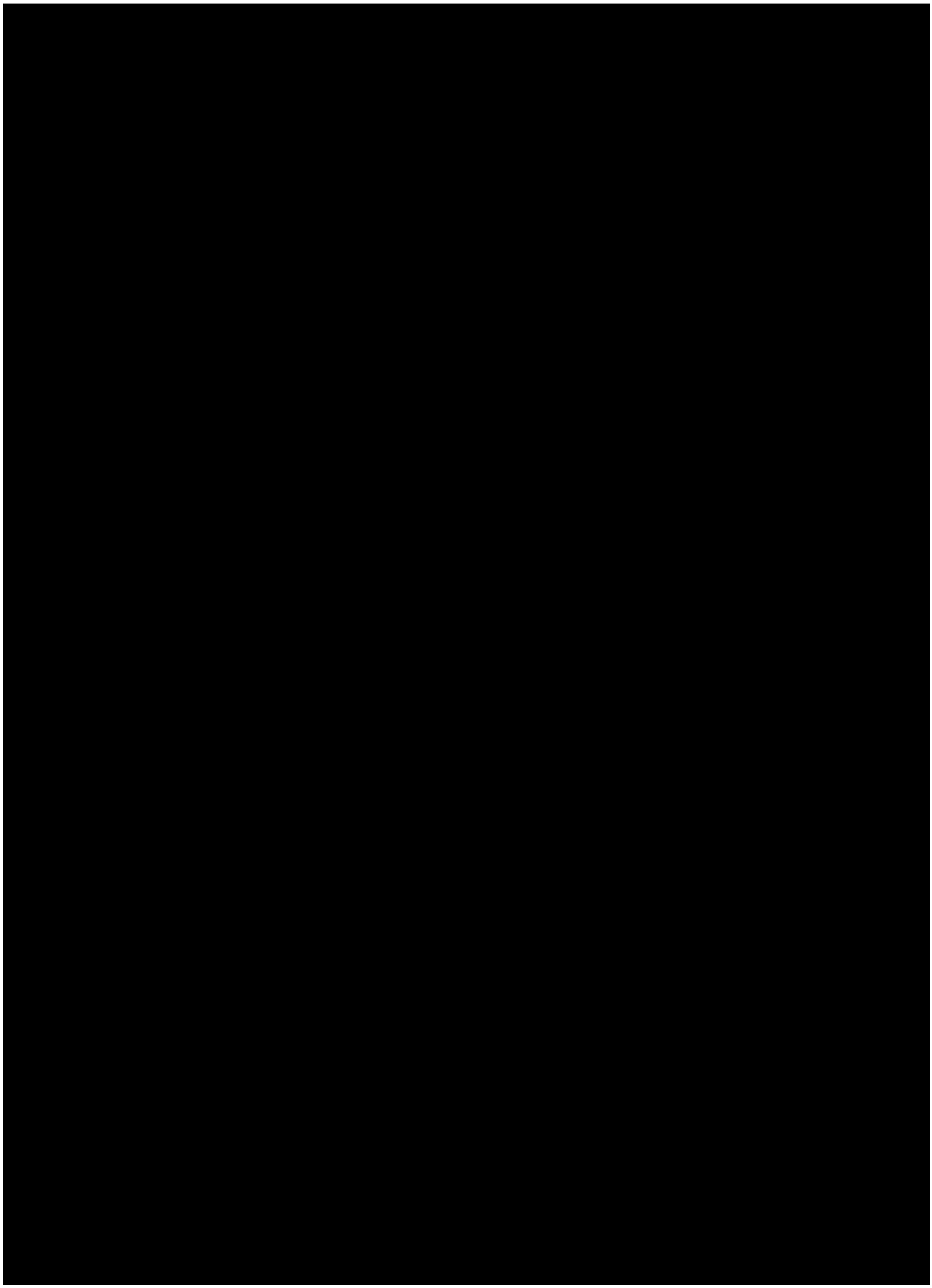
- Mean change from baseline to Week 18 in the following *European Organization for Research and Treatment on Cancer (EORTC) Quality of Life Core Questionnaire 30 items (QLQ-C30)* scores:
    - Physical functioning (PF)
    - Pain (PA)
    - Fatigue (FA)
    - Global health status / QoL (QL)
- Mean change from baseline to Week 6 of all the *EORTC QLQ-C30 PF, FA, PA, and QL scores* listed above
- Mean longitudinal change from baseline to Week 18 in the following scores obtained using items from the *EORTC QLQ-C30 and EORTC Item Library*:
    - Fatigue symptoms (FAS)
    - Fatigability (FAB)
    - Fatigue impact (FAI)
    - Pain descriptors (PAD)
    - Pain impact (PAI)
- Mean longitudinal change from baseline to Week 6 for each of the *EORTC Item Library FAS, FAB, FAI, PAD, and PAI scores* listed above

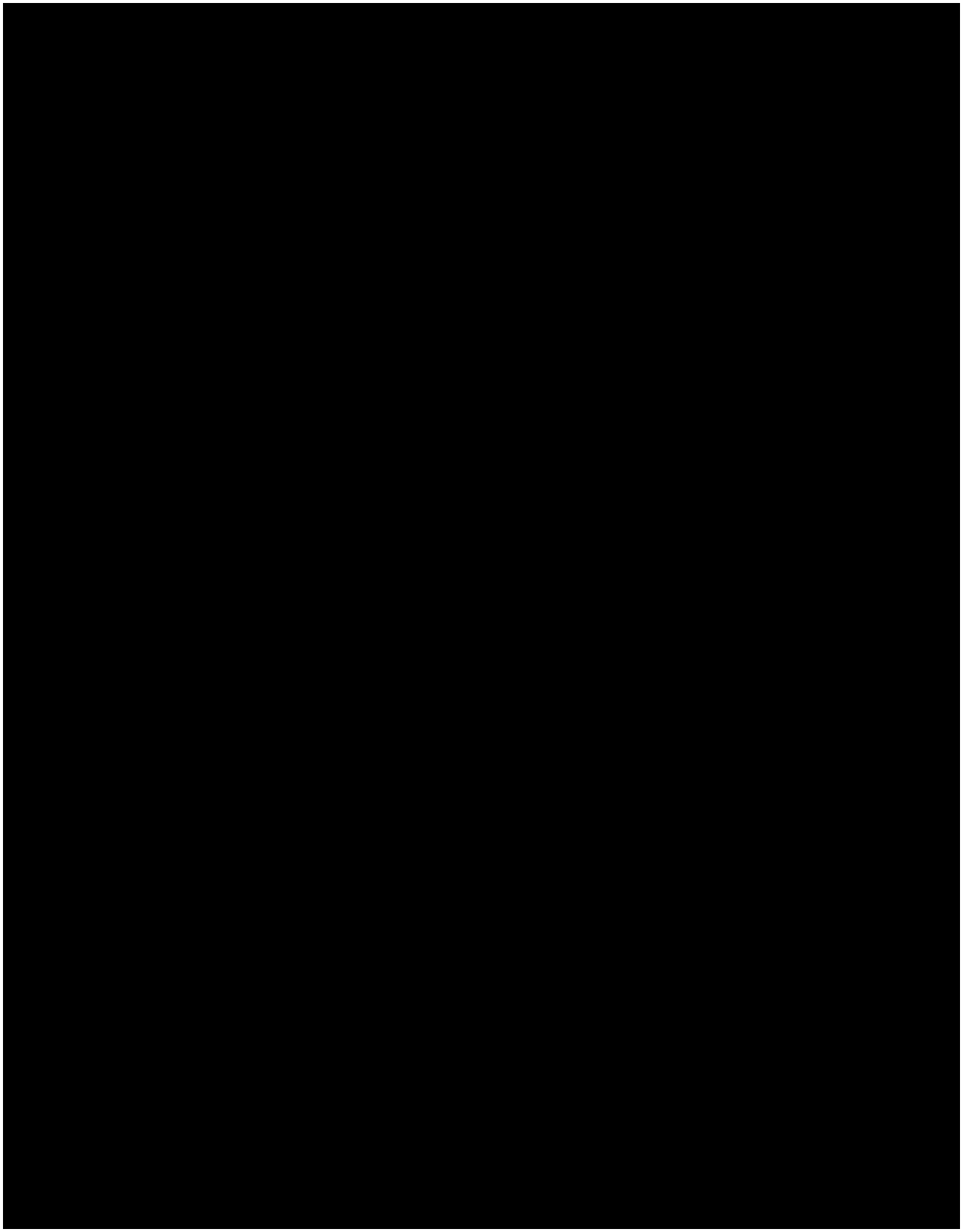
Safety endpoints

The secondary safety endpoints are as follows:

- Occurrence of treatment-emergent adverse events (AEs)
- Occurrence of treatment-emergent AEs leading to study drug discontinuation







6. GENERAL ANALYSIS DEFINITIONS

6.1 TREATMENTS

For reporting purposes, all randomized patients will be classified into one of the following:

- Brigimadlin 45 mg q3w
- Brigimadlin 30 mg q3w
- Doxorubicin

For efficacy analyses, patients will be analyzed as randomized.

For safety analyses, patients will be analyzed as treated.

The following trial periods are defined:

- The screening period is defined as the date of informed consent until the day prior to the date of first administration of trial medication.
- The treatment period is defined as the date from first administration of trial medication until the date of last administration of trial medication.
- For patients in the doxorubicin arm who receive subsequent brigimadlin in the trial, the cross-over treatment period is defined as the date from first administration of brigimadlin until the date of last administration of brigimadlin.
- The residual effect period (REP) is defined as the date of last administration of trial medication until a duration of 30 days. Patients in the doxorubicin arm who receive subsequent brigimadlin in the trial will have two REPs.
- The follow-up period is defined as the time after the on-treatment period until trial completion.

For safety summaries, data recorded from the date of the first administration of trial medication to 30 days after last administration of trial medication (i.e., including the REP) will be considered as on-treatment. For cross-over patients who have their first administration of subsequent brigimadlin prior to the end of their REP after last administration of doxorubicin, safety summaries will consider their on-treatment period on doxorubicin to be from date of first administration of doxorubicin until the date of first administration of subsequent brigimadlin in the trial.

Patients for whom randomized treatment group assignment has not been followed (i.e., cases where a patient's actual treatment is different from their randomized treatment assignment) will be handled on a case-by-case basis to be agreed upon at report planning meetings prior to database lock.

6.2 IMPORTANT PROTOCOL DEVIATIONS

Although all treated patients will be included in the safety analyses and all randomized patients will be included in the efficacy analyses, patients with important protocol deviations (iPDs) will be documented.

Definition and handling of iPDs in analyses is described in the iPD specification document stored in the trial master file (TMF) in the electronic Document Management System (eDMS). Identified iPDs will be finalized and documented before the (interim) database locks for reporting.

6.3 PATIENT SETS ANALYSED

The following patient sets are defined:

- Screened set:

  This patient set includes all patients who have signed the informed consent. The screened set will be used for patient disposition tables.

- Randomized set:

  This patient set includes all randomized patients, regardless of whether they have received any trial medication or not. Patients are assigned to receive brigimadlin or doxorubicin as randomized. The randomized set will be used for efficacy analyses and for HRQoL analyses. This patient set reflects the intention-to-treat (ITT) population.

- Treated set:

  This patient set includes all patients who were documented to have taken at least one dose of trial medication (brigimadlin or doxorubicin). The treated set will be used for safety analyses.

- Pharmacokinetic set:

  This patient set includes all patients who were documented to have received at least one dose of brigimadlin and have at least one observation for one PK endpoint without important protocol deviations relevant to the evaluation of PK. The pharmacokinetic set is used for the descriptive PK analyses.

- Cross-over set:

  This patient set includes all patients in the doxorubicin arm who receive subsequent brigimadlin in the trial. Patients will be analyzed as treated after cross-over.

6.5 POOLING OF CENTRES

This section is not applicable because centre/country is not included in the statistical model.

6.6 HANDLING OF MISSING DATA AND OUTLIERS

Missing or incomplete AE dates are imputed according to BI standards.

Censoring rules for PFS are in Table 7.3: 1 of the CTP. Censoring rules for OS are in Table 5.2.2: 2. Censoring rules for DOR are in Table 5.2.2: 3.

In general, missing data will not be imputed unless required for the following analyses and definitions, where the rules described below apply:

1) **Change of laboratory values from baseline**

   For missing laboratory data at Cycle 1 Day 1 (before the very first administration of any trial medication), the data of preceding visits will be used if available.

2) **Definition of on-treatment period and actual treatment**

   All reasonable efforts should be taken during the trial to obtain the dates of permanent discontinuation of trial medication. However, if the date of the very last administration is missing, then this date will be imputed with:

   - The last day of the month if only the month and year are given
   - The 31st of December of the year if only the year is given.

   If the imputed date leads to a date that is later than the date of the End of Treatment (EOT) visit, then the imputed date is the date of the EOT visit. If the imputed date leads to a date that is later than the death date, then the imputed date is the date of death.

3) **Randomization and stratification**

   In general, the data as reported in the eCRF will be used for analyses. If data (such as data relating to the assignment of stratification factors) are missing in the eCRF, then the data will be imputed with the values entered into the Interactive Response Technology (IRT) system. (Note that entries such as assignment of stratification factors are automatically transferred to the eCRF.) If the date of randomization differs between the IRT system and the eCRF, then the randomization date entered into the IRT system will be used.

4) **Partial death dates**

   If a partial death date is reported, then the following imputations will be performed. If the month is missing, then the month will be imputed with the month of January. If the day is missing, then the day will be imputed with day 1 (i.e., the first day of the month).

5) **Partial or missing start date of subsequent anti-cancer therapy or subsequent radiotherapy**

   If a partial start date of subsequent anti-cancer therapy or subsequent radiotherapy is reported, then the following imputations will be performed. If the month is missing, then the month will be imputed with the month of January. If the day is missing, then the day will be imputed with day 1 (i.e., the first day of the month). However if this leads to an imputed date on or before the stop date of trial medication, or if the start date is completely missing, then the start date will be imputed by the trial medication stop date + 1 day. Additionally, if death date is available, then the imputed start dates should be before the death date.

6) **Partial or missing stop date of subsequent anti-cancer therapy or subsequent radiotherapy**

   Unless otherwise specified, missing stop dates of subsequent anti-cancer therapies or subsequent radiotherapies will not be imputed. For the purpose of duration calculations,

partial stop dates will be imputed with the last day of the month (or year) up to the earlier of death, end of study, or data cut-off.

[redacted]

8) Partial concomitant medication start or end dates

In case of (partially) missing start and end dates of concomitant medication, the dates will be imputed to enable subsequent calculation (but not for display) by the "worst case" approach so that the extent of exposure to the concomitant therapy is maximal, i.e., the first day (month) of the month (year) for incomplete start dates and the last day (month) of the month (year) for incomplete end dates.

9) Partial first histological diagnosis date

If a partial first histological diagnosis date is reported, then the following imputations will be performed. If the month is missing, then the month will be imputed with the month of January. If the day is missing, then the day will be imputed with day 1 (i.e., the first day of the month).

10) PRO analyses

Missing items will be handled in the scoring of the PRO instrument as recommended by the developers of the PRO instrument (i.e., EORTC for EORTC QLQ-C30 and EuroQoL group for EQ-5D-5L). No imputation will be carried out for missing PRO scores.

11) PK data

Handling of missing PK data will be performed according to BI standard procedures as described in (4).

6.7 BASELINE, TIME WINDOWS AND CALCULATED VISITS

Study days and visits will be labelled according to the Flowcharts in the CTP.

Unless otherwise specified, baseline is defined as the time point closest to and prior to the first administration of trial medication.

- For some trial procedures (e.g., measurement of body weight), no time is specified, and the date is the same as the first administration date. In such cases, these measurements will be considered baseline.
- If there is no measurement earlier than the first administration of trial medication, then no baseline will be derived.
- For patients randomized but not treated, the latest screening value available will be used.

1) Laboratory values

Baseline is defined as the latest time point before the very first administration of trial medication. For laboratory assessments, time is recorded in addition to examination date; therefore, examination time must also be taken into account when defining baseline. A laboratory value on the same date as the first administration of trial medication is considered to be a baseline value if and only if the time of the laboratory value is before or the same as the time of first administration of trial medication.

If any of these times are missing, and the date of the laboratory assessment is equal to the date of first administration of trial medication, then the laboratory assessment will be considered as according to protocol, i.e., as prior to first administration of trial medication.

2) Cross-over set

For patients who cross-over from the doxorubicin arm to receive brigimadlin in the trial, baseline for the treatment period of brigimadlin is defined as the last value before the first administration of brigimadlin and after the end of doxorubicin treatment.

3) Imaging time windows

To identify whether consecutive imaging time points are missing for a given patient, a nominal time point (e.g., 6 weeks, 12 weeks, 18 weeks, 24 weeks, etc.) will be assigned to each image. This is achieved by creating windows for every radiological disease assessment. The windows are defined in Table 6.7: 1.

Table 6.7: 1 Nominal time points and windows for imaging

Nominal time point [weeks from Cycle 1 Day 1]	**Due date of scans [days from Cycle 1 Day 1]**	**Window [days]**
0 weeks	1 day	≤ 1 day
6 weeks	43 days	2 to ≤ 64 days
12 weeks	85 days	65 to ≤ 106 days
18 weeks	127 days	107 to ≤ 148 days
24 weeks	169 days	149 to ≤ 190 days
30 weeks	211 days	191 to ≤ 232 days
36 weeks	253 days	233 to ≤ 295 days
From Week 36 onwards, 12-week interval		
48 weeks	337 days	296 to ≤ 379 days
60 weeks	421 days	380 to ≤ 463 days
etc.		

For randomized patients who do not receive any trial medication, the date of randomization is the reference date for the nominal time points and windows for imaging instead of Cycle 1 Day 1.

If a patient does not have an image in one of the windows described in Table 6.7: 1, then the patient will be said to have "missed an assessment" for that time point.

4) PRO time windows

To identify whether consecutive PRO time points are missing for a given patient, a nominal time point will be assigned to each PRO assessment. This is achieved by creating windows for every PRO assessment. The windows are defined in Table 6.7: 2.

For randomized patients who do not receive any trial medication, the date of randomization is the reference date for the nominal time points and windows for PROs instead of Cycle 1 Day 1.

Table 6.7: 2 Nominal time points and windows for PROs

Nominal time point **[weeks from Cycle 1 Day 1]**	**Due date of PRO** **[days from Cycle 1 Day 1]**	**Window** **[days]**
0 weeks	1 day	≤ 1 day
QLQ C30, pain, fatigue *From Cycle 1 Day 1 until Week 48 or PD, whichever occurs first, every 3 weeks; and from PD (or Week 48) onwards annually*		
3 weeks	22 days	2 to ≤ 33 days
6 weeks	43 days	34 to ≤ 54 days
9 weeks	64 days	55 to ≤ 75 days
. . .	. . .	. . .
42 weeks	295 days	286 to ≤ 306 days
45 weeks	316 days	307 to ≤ 327 days
48 weeks	337 days	328 to ≤ 519 days
From PD or Week 48 onwards, 1-year interval		
100 weeks	701 days	520 to ≤ 883 days
152 weeks	1065 days	884 to ≤ 1247 days
etc.		
EQ5D, PGIC, PGIS *At Cycle 1 Day 1; at Week 15; from Week 15 to Week 48 or PD, whichever occurs first, every 12 weeks; and from PD (or Week 48) onwards annually*		
15 weeks	106 days	2 to ≤ 148 days
27 weeks	190 days	149 to ≤ 232 days
39 weeks	274 days	233 to ≤ 306 days
48 weeks	337 days	307 to ≤ 519 days

Nominal time point **[weeks from Cycle 1 Day 1]**	**Due date of PRO** **[days from Cycle 1 Day 1]**	**Window** **[days]**
From PD or Week 48 onwards, 1-year interval		
100 weeks	701 days	520 to ≤ 883 days
152 weeks	1065 days	884 to ≤ 1247 days
etc.		
PRO-CTCAE, FACT GP5 *From Cycle 1 Day 1 until Week 15 weekly; from Week 15 to Week 48 or PD, whichever occurs first, every 3 weeks; no assessments after PD (or Week 48)*		
1 week	8 days	2 to ≤ 12 days
2 weeks	15 days	13 to ≤ 19 days
3 weeks	22 days	20 to ≤ 26 days
. . .	. . .	. . .
15 weeks	106 days	104 to ≤ 117 days
18 weeks	127 days	118 to ≤ 138 days
21 weeks	148 days	139 to ≤ 159 days
. . .	. . .	. . .
42 weeks	295 days	286 to ≤ 306 days
45 weeks	316 days	307 to ≤ 327 days
48 weeks	337 days	≥ 328 days

7. PLANNED ANALYSIS

The following standards for End-Of-Text tables are defined:

- The set of summary statistics for continuous data is: N / Mean / Standard Deviation (SD) / Min / Median / Max. For tables that are provided for endpoints with some extreme data, median, quartiles, and percentiles should be preferred to mean, standard deviation, minimum, and maximum. If not otherwise specified, the abbreviation Pxx should be used for displaying the xx^{th} percentile. Other than the Min and Max, all statistics will be presented to one more decimal place than the raw data.

- Tabulations of frequencies for categorical data will include all possible categories and will display the number of observations in a category as well as the percentage (%) relative to the respective treatment group (unless otherwise specified, all patients in the respective patient set whether they have non-missing values or not). The precision for percentages will be one decimal place. The category "missing" will be displayed only if there are actually missing values.

- Time-to-event endpoints will be summarized in months to one decimal place. Hazard ratios (HRs), odds ratios, and corresponding confidence intervals (CIs) will be presented to two decimal places. P-values will be presented to four decimal places. All p-values will be one-sided if not specified otherwise. For HRs, if HR < 1, then the one-sided p-value = the two-sided p-value ÷ 2; otherwise, the one-sided p-value = 1 – the two-sided p-value ÷ 2. For odds ratios, if the odds ratio > 1, then the one-sided p-value = the two-sided p-value ÷ 2; otherwise, the one-sided p-value = 1 – the two-sided p-value ÷ 2.

Where applicable, conversion from days to weeks, months, and years will be as follows:

- Weeks = days ÷ 7
- Months = (days × 12) ÷ 365.25
- Years = days ÷ 365.25.

The interim CTR for the primary analysis of the primary endpoint PFS (based on blinded central independent review) in Phase III is called "the interim CTR" from here on.

The final CTR for the primary analysis of the selected secondary endpoint OS at the end of Phase III is called "the final CTR" from here on.

For all analyses (e.g., stratified analyses) using the stratification factor extent of disease (locally advanced vs. metastatic disease), in case the incorrect value was selected in the IRT, the actual value based on the data entered on the oncology history eCRF page will be used.

7.1 DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Only descriptive statistics on the randomized set are planned for this section of the interim CTR and the final CTR. In addition, descriptive statistics or listings on the cross-over set are planned for this section of the interim CTR and the final CTR, depending on the number of patients in the cross-over set at the time of analysis.

7.2 CONCOMITANT DISEASES AND MEDICATION

Only descriptive statistics on the randomized set are planned for this section of the interim CTR and the final CTR. In addition, descriptive statistics or listings on the cross-over set are planned for this section of the interim CTR and the final CTR, depending on the number of patients in the cross-over set at the time of analysis.

Concomitant diseases will be coded using the most current version of Medical Dictionary for Regulatory Activities (MedDRA).

Concomitant medications will be coded using the most current version of World Health Organization Drug Dictionary (WHO DD). Concomitant medications will be classified according to the Anatomical Therapeutic Chemical (ATC) classification system. The third ATC level will be used to categorize concomitant therapies by therapy type. In situations where a medical product may be used for more than one equally important indication, there are often several classification alternatives. As appropriate, patients receiving concomitant medications with more than one possible ATC level 3 category will be counted more than once. Explanatory footnotes will clarify the possible double counting.

Separate summaries will be produced of: (1) Previous and concomitant medications started prior to first administration of trial medication, and (2) concomitant medications started after first administration of trial medication and before the end of the REP.

7.3 TREATMENT COMPLIANCE

The amount of and time for which brigimadlin and doxorubicin is taken will be interpreted in light of treatment exposure, efficacy, and safety. Refer to Section 7.7 for further details on exposure analysis.

7.4 PRIMARY ENDPOINT

7.4.1 Primary analysis of the primary endpoint

The primary endpoint of PFS (based on blinded central independent review) will be analyzed using the randomized set of patients. The primary endpoint will be assessed and compared between the selected investigational arm and the doxorubicin control arm at an interim futility analysis at approximately the same time as the end of Phase II and at a primary analysis of PFS during Phase III. The analysis will be performed using the weighted inverse normal method described by Lehmacher and Wassmer [R14-1197] to control the familywise error rate for the adaptive design of the trial. More details are provided in Section 7 and Appendix 10.2 of the CTP.

Interim futility analysis

The interim futility analysis will be performed once the required number of approximately 56 PFS events occur from the selected investigational arm and the control arm.

PFS will be analyzed using a stratified log-rank test (locally advanced vs. metastatic) to assess the treatment effect. A technical one-sided alpha of 0.0005 will be spent at the interim futility analysis (and the rest of the one-sided alpha of 0.0245 will be spent at the primary analysis).

A non-binding futility boundary of HR = 0.80 for PFS is planned.

There is no plan to test for or claim superiority at the interim futility analysis.

A stratified Cox proportional hazards model will be used to estimate the HR of the treatment effect in the interim futility analysis and the corresponding asymptotic two-sided 95% Wald CI. HRs less than one will favor the selected brigimadlin arm. Breslow's method for handling ties will be used.

If the selected investigational arm passes the futility boundary, then the trial will continue to full enrollment.

Primary analysis of the primary endpoint PFS (based on blinded central independent review)

The primary analysis of PFS (based on blinded central independent review) will take place during Phase III at the time point when:

- Approximately 120 PFS events from the selected investigational arm and the control arm have occurred in patients enrolled in Stage 1 (Phase II);

  And

- Approximately 65 PFS events from the selected investigational arm and the control arm have occurred in patients enrolled in Stage 2 (Phase III).

Stage 1 (Phase II) is defined as the time period until 180 patients have been enrolled between the selected investigational arm and the control arm, and **Stage 2** (Phase III) is defined as the time period after that (when approximately 120 additional patients have been enrolled between the selected investigational arm and the control arm).

This approach follows option A (patient separation approach) in Jenkins *et al.* [R15-5218].

The primary PFS analysis will occur when the **pre-specified number of events from both stages** occur. The follow up time of Stage 1 patients is pre-specified until approximately 120 PFS events are observed from the selected investigational arm and the control arm. The follow up time of Stage 2 patients is planned until approximately 65 PFS events are observed from the selected investigational arm and the control arm.

The primary endpoint will be analyzed using a stratified (locally advanced vs. metastatic) log-rank test to assess the treatment effect. Closed testing is implemented using the Dunnett test to control the Type I error rate due to the dose selection in the trial. For the patients enrolled in Stage 1, the p-value is based on the one-sided Dunnett test adjusted p-value for dose selection derived from the stratified log-rank test statistics for PFS. For the patients enrolled in Stage 2, the p-value is based on the one-sided p-value derived from the stratified log-rank test statistics for PFS. The test for statistical significance at the primary PFS analysis will then be based on combining these p-values via the weighted inverse combination function. This process is based on independent, normally distributed stratified log-rank test statistics, which ensures control of the Type I error rate. Additional technical details of the combination test approach can be found in Appendix 10.2 of the CTP.

To estimate the HR for PFS at the primary PFS analysis during Phase III, the median unbiased estimator [R14-2334] will be used as the primary estimator. HRs less than one will favor the selected brigimadlin arm. The 95% confidence interval for the HR will then be

calculated as a repeated confidence interval (RCI) according to [R15-0928]. In addition, as a secondary estimator, the partial maximum likelihood estimator from the stratified Cox proportional hazards model (stratified by locally advanced vs. metastatic) will be used.

LR1: The stratified log-rank statistic between the selected investigational arm and the control arm derived from patients enrolled in Stage 1 until the planned number of events d_1 is observed.

LR2: The stratified log-rank statistic between the selected investigational arm and the control arm derived from patients enrolled in Stage 2 until the planned number of events d_2 is observed.

$$\widehat{\theta_1} = LR1\sqrt{\frac{4}{d_1}} \sim N\left(\theta, \frac{4}{d_1}\right)$$

$$\widehat{\theta_2} = LR2\sqrt{\frac{4}{d_2}} \sim N\left(\theta, \frac{4}{d_2}\right)$$

where θ = log(HR) and HR is the assumed Hazard ratio; $\widehat{\theta_1}$ is the estimation of θ based on data from Stage 1, and $\widehat{\theta_2}$ is the estimation of θ based on data from Stage 2.

According to [R14-2334], the median unbiased estimator of θ is given by:

$$\widehat{\theta_m} = \frac{w_1\sqrt{d_1}\widehat{\theta_1} + w_2\sqrt{d_2}\widehat{\theta_2}}{w_1\sqrt{d_1} + w_2\sqrt{d_2}} = \frac{w_1\sqrt{d_1}LR1\sqrt{\frac{4}{d_1}} + w_2\sqrt{d_2}LR2\sqrt{\frac{4}{d_2}}}{w_1\sqrt{d_1} + w_2\sqrt{d_2}} = 2\frac{w_1LR1 + w_2LR2}{w_1\sqrt{d_1} + w_2\sqrt{d_2}}$$

where w_1 and w_2 are the pre-defined weights for Stages 1 and 2 specified in CTP Section 10.2. Thus, the median unbiased estimator for the HR is $\exp(\widehat{\theta_m}) = \exp(2\frac{w_1LR1 + w_2LR2}{w_1\sqrt{d_1} + w_2\sqrt{d_2}})$.

If the trial continues to Stage 2, then the two-sided 95% RCI for the HR is given by:

$$\exp\left(\widehat{\theta_m} \pm \frac{2z_2}{w_1\sqrt{d_1} + w_2\sqrt{d_2}}\right)$$

where z_2 is the critical value for rejection of the null hypothesis at Stage 2.

Kaplan-Meier estimates will be used to display the distribution of PFS for each treatment group on a Kaplan-Meier curve. To support the plot, estimated survival probabilities will be tabulated. In addition, the Kaplan-Meier estimates will be used to provide estimates of the median, 25th, and 75th percentiles.

Suitable methods will be used to check the assumption of proportional hazards.

If statistical significance is obtained for PFS at the primary PFS analysis during Phase III, then the selected secondary endpoints (i.e., ORR and OS) will be tested following a hierarchical framework. More details are in Section 7.5.2.

Dunnett test

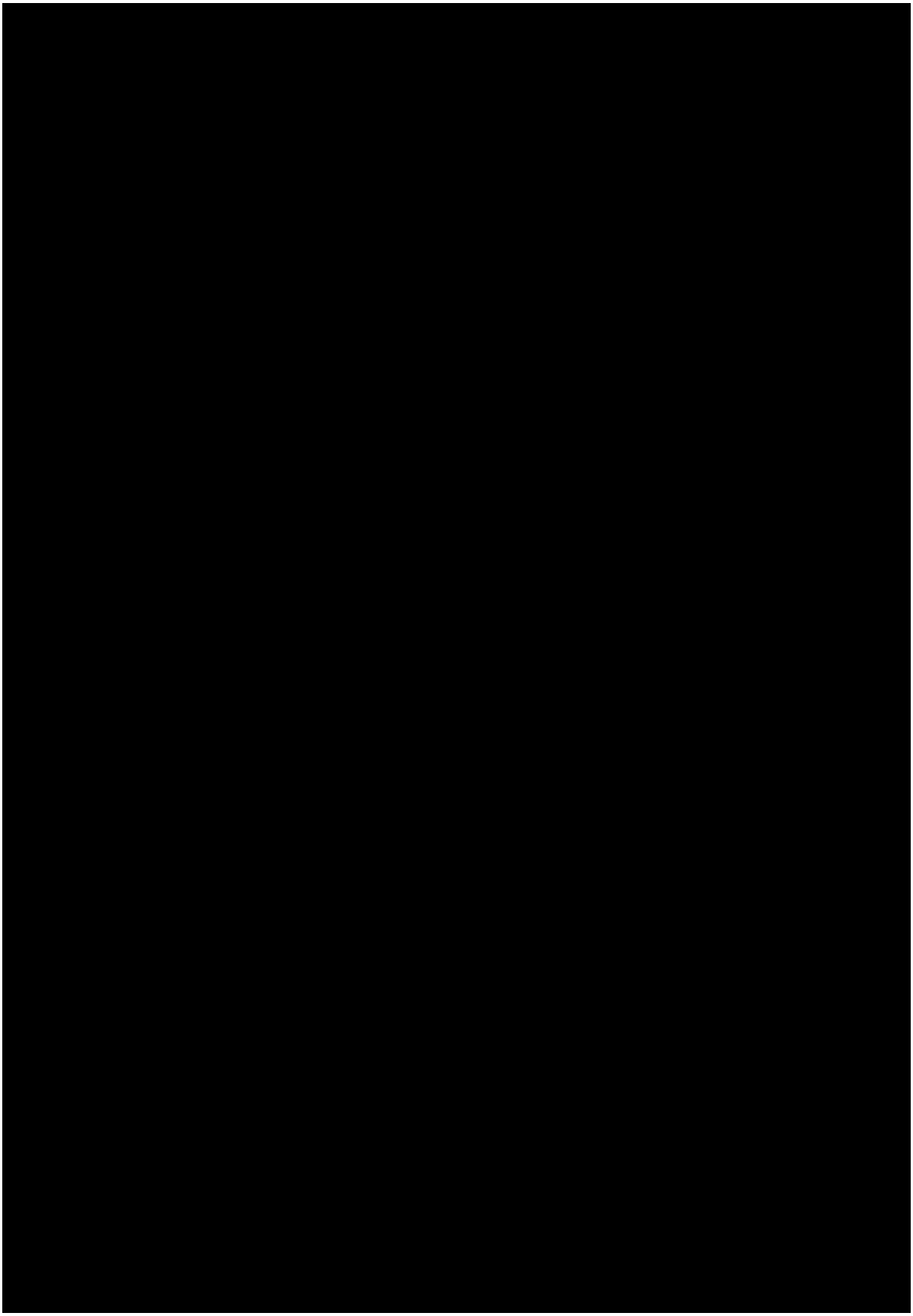
As stated in Appendix 10.2 of the CTP, the Dunnett test is a multiple comparison procedure for the comparison of multiple investigational treatments with a single control. Closed testing is implemented using the Dunnett test to control the Type I error rate due to the dose selection in the trial.

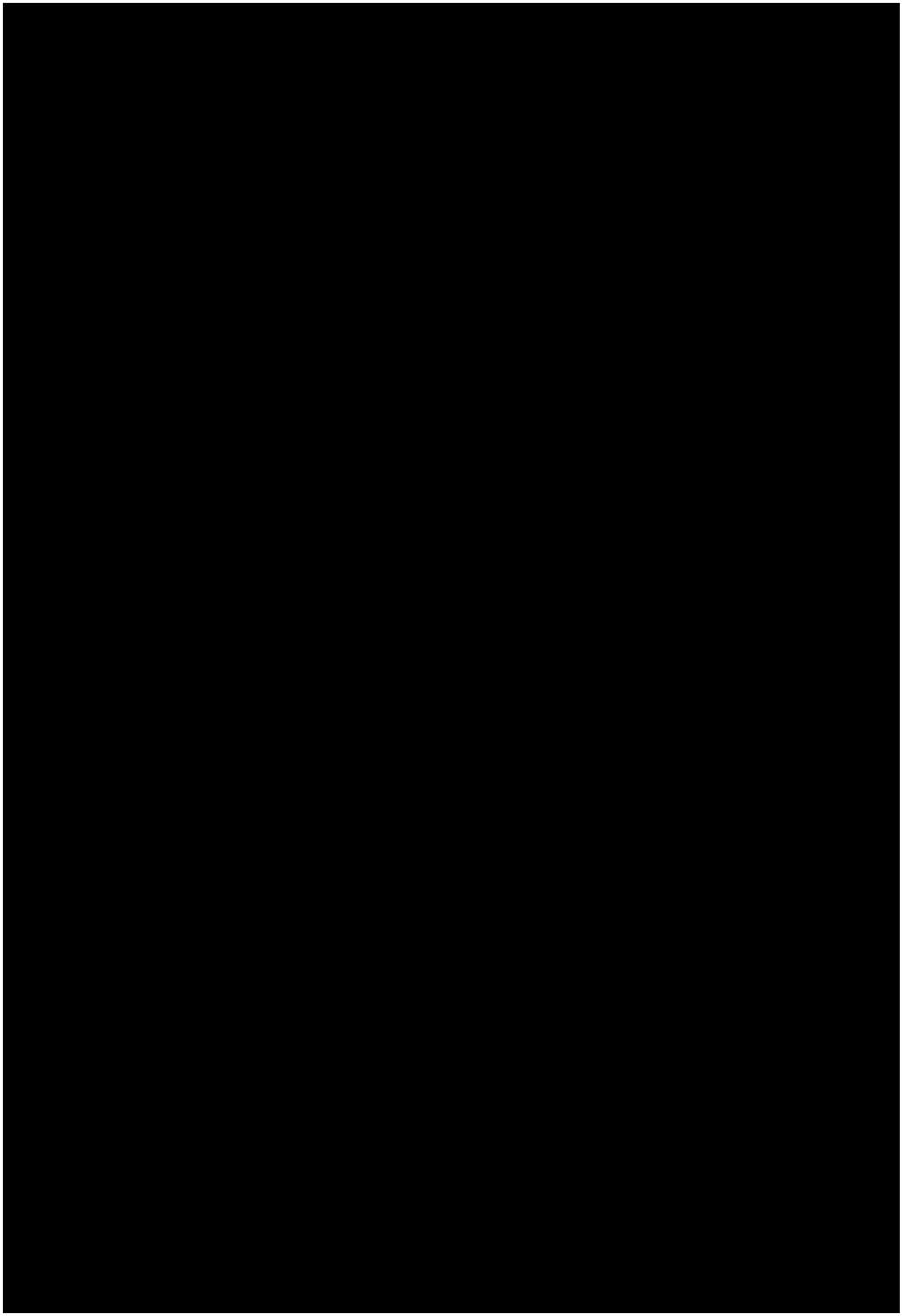
The Dunnett test will be implemented using the "rpact" R package. "rpact" is a statistical program module characterized as a comprehensive, validated R package that enables the simulation and analysis of confirmatory adaptive designs with continuous, binary, or survival endpoints. "rpact" can be downloaded per CRAN and is available as open-source software licensed under LGPL3. The "rpact" package followed a rigid validation procedure with documented testing and verification of its functionality. A comprehensive validation documentation (compliant to FDA/GxP guidelines) is available upon request. For more information, please see https://www.rpact.com/ and https://www.rpact.org/with Vignettes.

R code for Dunnett test:

```
require(rpact)

datasetSurvival <- getDataset(
  events1 = c(120),
  events2 = c(event_unselected),
  logRanks1 = c(LR_selected),
  logRanks2 = c(LR_unselected),
  allocationRatios1 = c(1),
  allocationRatios2 = c(ratio_unselected)
)

x <- getAnalysisResults(dataInput = datasetSurvival, intersectionTest =
"Dunnett", directionUpper = F)

adjust_pvalue <- max(x$.closedTestResults$adjustedStageWisePValues[1,],
x$.closedTestResults$adjustedStageWisePValues[2,])
```

In the code above, `event_unselected` is the total number of events between the unselected investigational arm and the control arm among patients enrolled in Stage 1. `LR_selected` is the stratified log-rank statistic between the selected investigational arm and the control arm, and `LR_unselected` is the stratified log-rank statistic between the unselected investigational arm and the control arm. It should be noted that `LR_selected` and `LR_unselected` are the "unsquared" version of the log-rank test statistics, which should be standard normally distributed under the null hypothesis. `ratio_unselected` is the number of patients randomized to the unselected investigational arm divided by the number of patients randomized to the control arm.





7.5 SECONDARY ENDPOINTS

7.5.1 Key secondary endpoint

This section is not applicable as no key secondary endpoint has been specified in the protocol.

7.5.2 (Other) Secondary endpoints

If statistical significance is obtained for PFS at the primary analysis of PFS during Phase III, then the selected secondary endpoints (i.e., ORR and OS) will be tested following a hierarchical testing framework. Due to the hierarchical testing procedure chosen, no alpha adjustment is required to account for multiple testing. If statistical significance is obtained for PFS at the primary analysis of PFS during Phase III, then ORR will be tested at one-sided alpha level 0.0245 during Phase III at the same time as the primary analysis of PFS (in the interim CTR). If statistical significance is obtained for both PFS and ORR, then the alpha will be carried over to the primary analysis and test of OS at the end of Phase III (in the final CTR). If statistical significance is obtained for PFS but not for ORR, then OS will continue to be followed up until the pre-specified number of OS events are met at the end of Phase III and then analyzed without being formally tested.

Objective response (OR) based on *blinded central independent review*

ORR gives the percentage of patients with objective tumor response. If statistical significance is obtained for PFS at the primary analysis of PFS, then the Cochran-Mantel-Haenszel test

will be used to test for a difference between the selected investigational treatment arm and the doxorubicin control arm for ORR, and ORR will be tested at a one-sided alpha level of 0.0245 at the same time as the primary analysis of PFS in the interim CTR. The stratification factor will be adjusted for in the analysis (locally advanced vs. metastatic).

The two-sided p-value will be generated from the Cochran-Mantel-Haenszel test statistic. The one-sided p-value will then be inferred from the two-sided p-value. The weighted inverse normal method described by Lehmacher and Wassmer [R14-1197] combining one-sided p-values from two stages will be the analysis method. The weights for ORR analysis are chosen based on the proportion of the planned numbers of patients randomized to the selected treatment arm and the control arm from Stage 1 and the planned number of patients randomized in Stage 2. The corresponding planned number of patients are 180 (60%) in Stage 1 and 120 (40%) in Stage 2. Therefore, the weights are $\sqrt{0.6}$ for Stage 1 and $\sqrt{0.4}$ for Stage 2.

The Stage 1 Dunnett's adjusted p-value will be calculated only based on patients enrolled at Stage 1, and the Stage 2 p-value will be calculated only based on patients enrolled at Stage 2. More details are in the statistical appendix of the CTP.

To estimate the odds ratio for ORR, the median unbiased estimator [R14-2334] will be used as the primary estimator. Since there are no sample size re-estimation, effectively the Cochran-Mantel-Haenszel method will be used to estimate the odds ratio and the 95% CI. In addition, as a secondary estimator, an odds ratio and corresponding 95% CI will be estimated by a logistic regression adjusting for extent of disease.

For the final CTR, ORR (based on blinded central independent review) will only be summarized descriptively for all three treatment groups in the randomized set. In addition, these descriptive analyses will be repeated for the subgroup variables for efficacy detailed in Section 6.4 for the interim and final CTRs.

Overall survival (OS)

For the interim CTR, no hypothesis testing for OS using the combination test approach will be repeated. The HR and Kaplan-Meier estimates will be analyzed. No subgroup analyses will be performed for the early, exploratory analysis of OS in the interim CTR.

If statistical significance is obtained for both PFS and ORR at the primary analysis of PFS, then the alpha of one-sided level 0.0245 will be carried over to the primary analysis and test of OS in the final CTR.

The primary analysis of OS will occur at the end of Phase III at the time point when:

- Approximately 133 OS events from the selected investigational arm and the control arm have occurred in patients enrolled in Stage 1 (Phase II),

  And

- Approximately 82 OS events from the selected investigational arm and the control arm have occurred in patients enrolled in Stage 2 (Phase III).

For the primary analysis of OS at the end of Phase III, the testing process will be similar to the analysis described for PFS (i.e., the Dunnett test and combination of Stage 1 and Stage 2 p-values via the weighted inverse combination function will be used for OS), and OS would be tested at a one-sided level of 0.0245.

Kaplan-Meier estimates of the survival function of OS, median OS, and the corresponding 95% CI using Greenwood variance that is incorporated into the Brookmeyer and Crowley method [R09-6372] with a loglog transformation, as well as the p-value of stratified log-rank test (stratified by locally advanced vs. metastatic) on OS, will be provided.

To estimate the HR for OS, the median unbiased estimator [R14-2334] will be used as the primary estimator. The 95% CI for the HR will then be calculated as repeated confidence interval (RCI) according to [R15-0928]. In addition, as a secondary estimator, the partial maximum likelihood estimator from the stratified Cox proportional hazards model (stratified by locally advanced vs. metastatic) will be used.

To address cross-over effects, the following sensitivity analysis of OS is planned. [redacted]

[redacted]

Duration of response (DOR) based on *blinded central independent review*

Duration of confirmed OR will be summarized by its median and quartiles derived using the Kaplan-Meier estimation procedure. DOR can only be calculated for patients that have a confirmed OR based on *blinded central independent review*.

Disease control (DC) based on *blinded central independent review*

Disease control rate gives the percentage of patients with DC. Logistic regression will be used to explore the difference between treatment arms for disease control rate. The stratification factor will be included as a covariate in the logistic regression model (locally advanced vs. metastatic). An odds ratio and corresponding 95% CI will be generated using the likelihood ratio CI. Odds ratios > 1 favor brigimadlin.

Health-Related Quality of Life (HRQoL)

All scores for the *EORTC QLQ-C30, EORTC Item Library* and *EQ-5D-5L* will be calculated according to published scoring algorithms.

The *EORTC QLQ-C30* is composed of both multi-item scales and single-item measures. These include five functional scales, three symptom scales, a global health status / QoL scale, and six single items. The scales are scored based on a linear transformation of the raw scores to a scale from 0 to 100. The raw scores are the averages of the item scores contributing to the respective scale.

The *EORTC Item Library* measures are composed of five multi-item scales. The scores for Fatigue Symptoms, Fatigability, Pain Descriptors, Fatigue Impact, and Pain Impact using items from the *EORTC QLQ-C30 and EORTC Item Library* will be calculated according to the same principles as the *EORTC QLQ-C30* symptom scale scores.

Questionnaire Scale	Scale	Number of items	Item range	Item Numbers	Function Scales
EORTC QLQ-C30					
Global health status/ QoL	QL	2	6	29, 30	
Physical Functioning	PF	5	3	1 to 5	F
Role Functioning	RF	2	3	6, 7	F
Emotional Functioning	EF	4	3	21 to 24	F
Cognitive Functioning	CF	2	3	20, 25	F
Social Functioning	SF	2	3	26, 27	F
Fatigue	FA	3	3	10, 12, 18	
Nausea and vomiting	NV	2	3	14, 15	
Pain	PA	2	3	9, 19	
Dyspnea	DY	1	3	8	
Insomnia	SL	1	3	11	
Appetite loss	AP	1	3	13	
Constipation	CO	1	3	16	
Diarrhea	DI	1	3	17	
Financial difficulties	FI	1	3	28	
EORTC Item Library Pain					
Pain descriptors	PAD	8	3	1 to 8	
Pain impact	PAI	10	3	9 to 18	
EORTC Item Library Fatigue					
Fatigue symptoms	FAS	8	3	1 to 8	
Fatigability	FAB	12	3	9 to 20	
Fatigue impact	FAI	5	3	21 to 25	

To calculate the scale scores, a raw score (RS) is first calculated for each scale as the mean of the items of the scale.

$$RS = \frac{(I_1 + I_2 + \cdots + I_n)}{n}$$

Then, the Functional scale scores are transformed on a 0-100 scale as follows:

$$Score = \left(1 - \frac{(RS - 1)}{range}\right) \times 100$$

and for Symptom scales/ items and Global health status/ QoL:

$$Score = \left(\frac{(RS - 1)}{range}\right) \times 100$$

A scale score is calculated if at least 50% of the items of a scale are completed (i.e., non-missing). This calculation is from the EORTC Scoring Manual [R07-2064]. In the presence of missing items where at least 50% of items of a scale are completed, the raw score (RS) is computed as the average of answered items.

PRO compliance and completion

Evaluable and complete PRO data will be described per instrument as follows:

- Number of participants for whom a questionnaire is expected to be completed (i.e. participants who are still ongoing in the study and expected to complete the questionnaire as per the schedule of assessments);
- Number of participants with evaluable questionnaire (i.e., questionnaire with at least one non-missing item) by visit;
- Number of participants with all items on the questionnaire completed by visit;

Compliance and Completion will be calculated at each scheduled visit for the *EORTC QLQ-C30* and *EORTC Item Library* measures for fatigue and pain as follows:

$$\%\ \text{Completion} = 100 \times \frac{\text{number of participants with all EORTC PRO items available at the visit}}{\text{number of participants for whom EORTC PRO data is expected at the visit}}$$

$$\%\ \text{Compliance} = 100 \times \frac{\text{number of participants with evaluable EORTC PRO available at the visit}}{\text{number of participants for whom EORTC PRO data is expected at the visit}}$$

Similar completion rates will also be calculated at each visit for the *EQ-5D-5L*.

Descriptive analyses of scores over time

Descriptive statistics will be calculated across all visits for the actual values and change from baseline of *EORTC-QLQ-C30 PF, FA, PA, and QL scores*, as well as *EORTC Item Library FAS, FAB, FAI, PAD, and PAI scores*. Descriptive statistics will also be calculated across all visits for the actual and change from baseline of the *EQ-5D-5L VAS*. The distribution of *EQ-5D-5L* levels of perceived problems (1=no problem, 2=slight, 3=moderate, 4=severe, and 5=extreme problem) will be described for each dimension (mobility, self-care, usual activities, pain / discomfort, and anxiety / depression) across all visits as frequencies and percentages.

A graphical display showing the mean change in score from baseline at all visits in the investigational arm and in the doxorubicin arm will be created for each of *EORTC QLQ-C30 PF, FA, PA, and QL scores, EORTC Item Library FAS, FAB, FAI, PAD, and PAI scores*, and *EQ-5D-5L VAS*.

Descriptive analysis of other HEOR variables over time

Analgesic intake will be described using the Analgesic Quantification Algorithm (AQA). The number and percentage of participants in each AQA category will be calculated at each visit.

Longitudinal modelling of PRO scores

The change in the PRO scores assessing the key concepts of physical functioning, the key symptoms related to liposarcoma and their impact (fatigue, fatigability, pain, fatigue impact, and pain impact) and overall HRQoL will be analyzed using piecewise linear mixed-effects models with different slope of change over two periods: Baseline to Week 6 and after Week 6.

The *EORTC QLQ-C30 PF, FA, PA, and QL scores*, as well as the *EORTC Item Library FAS, FAB, FAI, PAD and PAI scores* will be analyzed using piecewise linear mixed models. A separate model will be run for each PRO scale score. All PRO assessments until Week 48 will be included in the model. Data after switch from brigimadlin to a subsequent anti-cancer therapy following PD and data after cross-over from doxorubicin to brigimadlin will not be included in the model. Data after switch from brigimadlin to a subsequent anti-cancer therapy reasons other than PD will be included in the model.. The response variable will be the change in the PRO score from baseline at each available post-baseline visit. Time since baseline will be included as a continuous variable as a fixed effect with a parametrization distinguishing two periods, before and after Week 6. The change in PRO scores will therefore be assumed to possibly follow different slopes before or after Week 6. The stratification factor (locally advanced vs. metastatic) will be included in the model as a covariate (fixed effect). The baseline PRO score will be included in the model as a covariate.

The model is summarized in the formula below:

$$Y_{ij} = \beta_0 X_{i0} + \beta_1 X_{i1} + \left(\beta_2 + \beta_3 t_{ij} + \beta_4 t_{ij}^{[6]}\right) + \left(\beta_5 + \beta_6 t_{ij} + \beta_7 t_{ij}^{[6]}\right) \cdot X_{i2} + (b_{i0} + b_{i1} t_{ij} + b_{i2} t_{ij}^{[6]}) + \varepsilon_{ij}$$

where Y_{ij} is the observed change from baseline of the PRO score for individual *i* at time *j*, X_{i0} is the baseline value, X_{i1} is the indicator for the stratification factor, X_{i2} is the indicator for the randomized treatment, t_{ij} is the actual time of assessment since baseline in study days, $t_{ij}^{[6]} = \max(t_{ij} - 43, 0)$ is the time in days since study day 43, i.e. the planned Week 6, and $\beta_0 \ldots \beta_7$ are the corresponding fixed effects.

Random effects will include a random intercept, b_{i0}, and two random slopes b_{i1} and b_{i2}, (one for each period, before or after Week 6). An unstructured covariance matrix will be used for the random effects. In case of nonconvergence or computational issues, an independent structure will be used instead.

Within-subject correlation will be induced by the random effects, an independent covariance structure will be used for the error term ε_{ij}. As the PRO assessments are equally spaced (every 3 weeks), a first-order autoregressive covariance matrix, AR(1), will be used if the model fails to converge. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom and adjust standard errors. All missing data will be assumed to be missing at random.

The adjusted longitudinal mean change in score from baseline to Week 18 (average of LS Means up to Week 18) and its associated 95% CI will be estimated and compared between the investigational arm and the doxorubicin arm. More specifically, the average over the LS Means from the model at Week 3, Week 6, Week 9, Week 12, Week 15 and Week 18 will be

taken and interpreted as the average longitudinal change from baseline up to Week 18. Additionally, it will also be estimated and compared at Week 6. In this case the average in LS Means is taken over Week 3 and Week 6.

The difference between the adjusted longitudinal mean change from baseline in *EORTC QLQ-C30 PF, FA, PA, and QL scores* between the investigational arm and the doxorubicin arm estimated by the piecewise linear mixed models will be interpreted using the published guidelines for interpretation of longitudinal difference in *QLQ-C30* scores [R14-1929]: differences in *EORTC QLQ-C30 PF, FA, and QL scores* between 5 and 10 points will be considered small and above 10 points medium; differences in *EORTC QLQ-C30* PA scores between 3 and 11 points will be considered small and above 11 points medium. For *EORTC Item Library FAS, FAB, FAI, PAD and PAI scores*, no guidance exists for their interpretation. The clinical meaningfulness of the difference in adjusted mean change for these scores will therefore initially be interpreted using similar benchmarks: differences in scores between 5 and 10 points will be considered small and above 10 points medium. The time to deterioration analyses will be using the medium change thresholds.

Table 7.5.2: 1 Estimands table for primary analysis of HRQoL

Estimand attribute	**Descriptions**
Population	Patients with advanced dedifferentiated liposarcoma
Treatment condition(s)	Brigimadlin vs. doxorubicin as first line
Variable (outcome)	PRO scores assessing the key concepts of physical functioning, the key symptoms related to liposarcoma and their impact (fatigue, fatigability, pain, fatigue impact, and pain impact) and overall HRQoL at Week 6 and Week 18: *EORTC QLQ-C30 PF, FA, PA, and QL scores*, as well as *EORTC Item Library FAS, FAB, FAI, PAD and PAI scores*.
Handling of intercurrent events	- *Start of subsequent therapy following PD*: Hypothetical strategy – Data after start of subsequent therapy following PD will be censored or excluded. The statistical model will estimate the treatment effect if patients had not taken subsequent therapy. - *Start of subsequent therapy for other reasons*: Treatment Policy strategy – Data after the switch to subsequent anti-cancer therapy will be included in the statistical models. - *Treatment crossover:* Hypothetical strategy – Crossover patients from the doxorubicin to brigimadlin arm will be handled by a hypothetical strategy. Data after crossover will be censored or excluded. - *Treatment discontinuation*: Treatment Policy strategy – Data after the treatment discontinuation will be included in the statistical models.

	- *Death*: Hypothetical strategy – The statistical models will treat the missing data arising from death as missing-at-random and estimate the treatment effect in the scenario where the patient had not died. - *Opioid use (for EORTC QLQ-C30 PA* and *EORTC Item Library PAD and PAI scores only)*: Treatment policy strategy – Opioid use measured as any pain assessment obtained at a visit when the participant has an AQA score of 2 or greater; data after opioid use will be included in the statistical models to estimate the treatment effect of brigimadlin or doxorubicin in combination with any pain medication, including opioids, on patients.
Summary measure	Difference between brigimadlin vs. doxorubicin in mean change from baseline to Week 18 (resp. Week 6)

Time to deterioration analysis

A time to first deterioration (TTD) analysis will be performed in the randomized set of the full Phase II/III data for the PRO scores of interest (*EORTC QLQ-C30 PF, FA, PA, and QL scores*, plus *EORTC Item Library FAS, FAB, FAI, PAD, and PAI scores*).

TTD will be defined as the time between the randomization date and the occurrence of the first meaningful within-patient deterioration. For each PRO score a meaningful within-patient deterioration is defined as:

- *EORTC QLQ-C30 PF and QL scores*: A change from baseline < -10 points.
- *EORTC QLQ-C30 FA score* and *EORTC Item Library FAS, FAB, FAI, PAD and PAI scores*: A change from baseline > 10 points.
- *EORTC QLQ-C30 PA score*: A change from baseline > 11 points.

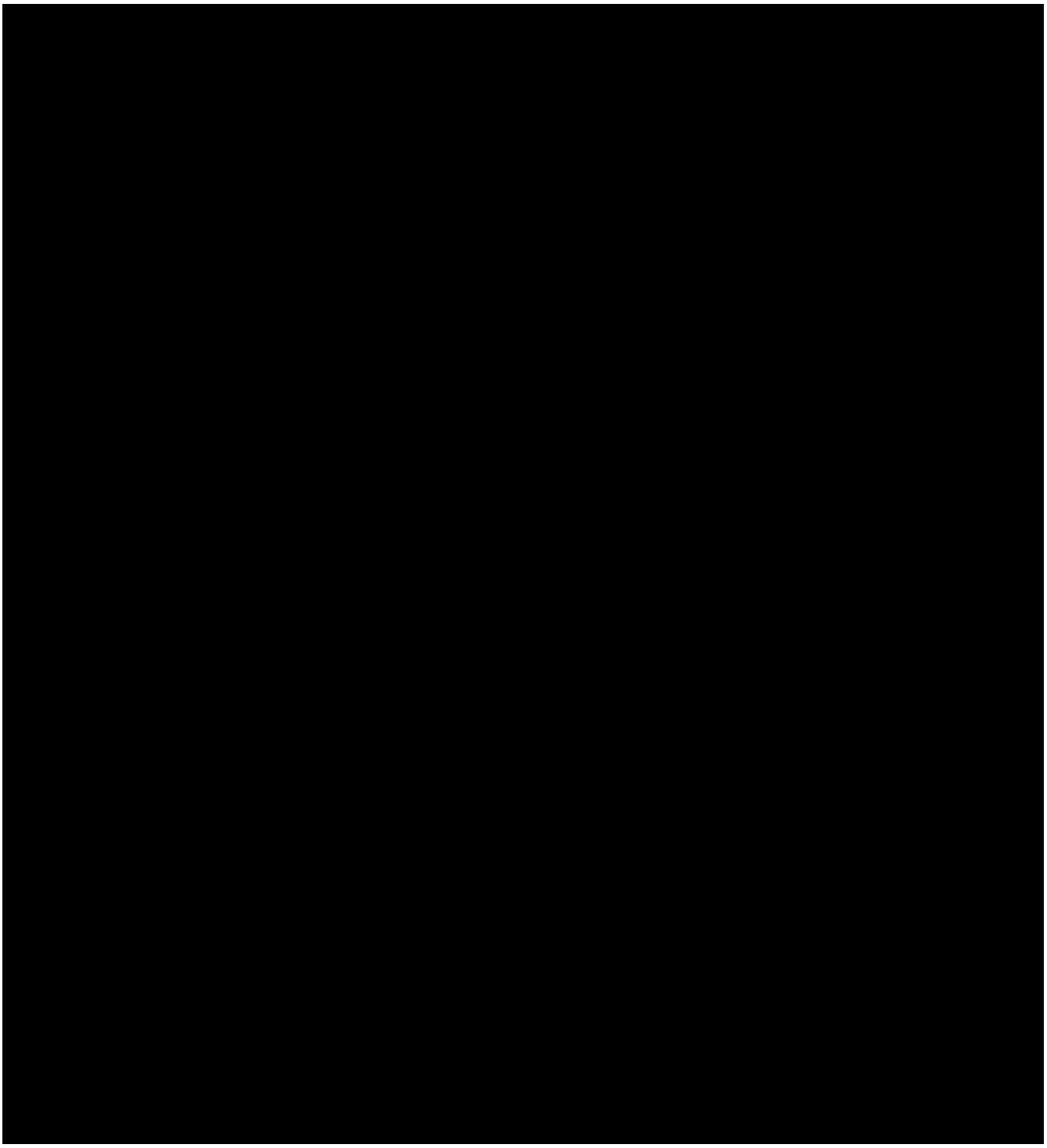
Kaplan-Meier estimates of the time to first deterioration function for each PRO score will be produced to compare the treatment groups, along with a summary of median deterioration free survival time, and deterioration free survival rates at weeks 6, and 18 and corresponding two-sided 95% confidence intervals. They will be accompanied by log-rank tests. Additionally, stratified Cox proportional hazard models will be used to further investigate TTD, by producing hazard ratio estimates and their corresponding two-sided 95% Wald confidence intervals. The stratification factor for the model will be the stratification of the study (locally advanced vs. metastatic). Breslow's method for handling ties will be used.

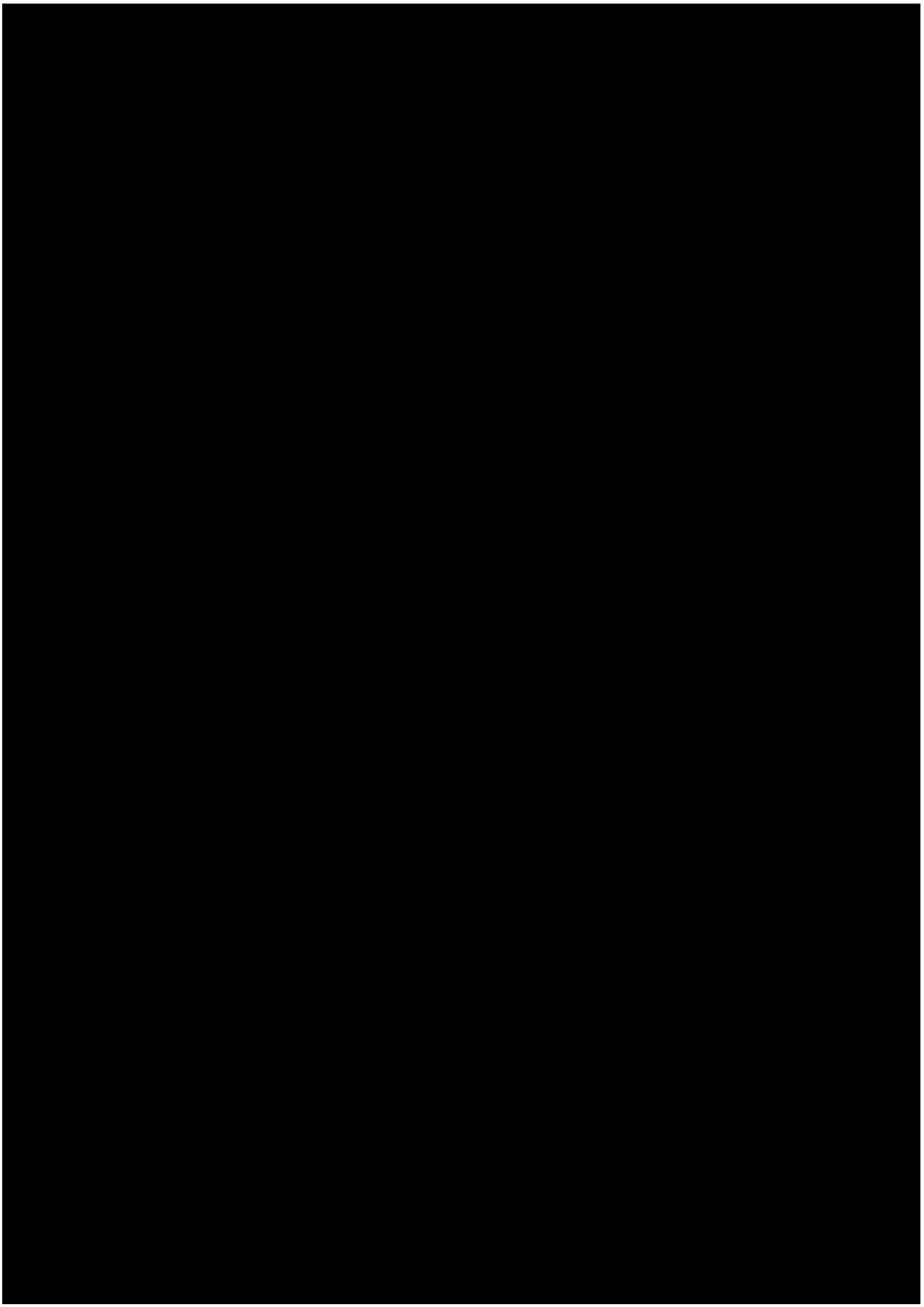
The following censoring rule will be applied in the analyses:

- Participants without deterioration will be censored at the time of the last available PRO assessment. Patients without deterioration before any switch to subsequent anti-cancer therapy following PD based on *blinded central independent review*, cross over from doxorubicin to brigimadlin, or death will be censored at last PRO assessment prior to such. Patients who were randomized but not treated, or did not have a baseline PRO measurement, will be censored at date of randomization.

Sensitivity and supplementary analyses

Further HRQoL analyses, including sensitivity and supplementary analyses for the main analysis described above, will be specified in a separate HSAP. The HSAP will also specify some psychometric analyses to generate supportive information on the measurement properties of the *EORTC Item Library FAS, FAB, FAI, PAD, and PAI scores*, as well as information on their meaningful change.





7.7 EXTENT OF EXPOSURE

All exposure analyses described below will be performed on the treated set and the cross-over set for the interim CTR and the final CTR.

Total number of cycles

A descriptive summary of the **total number of cycles** of brigimadlin or doxorubicin will be produced for both the treated set and the cross-over set.

Duration of treatment with trial medication (or cross-over treatment)

A descriptive summary of **total duration of treatment with trial medication**, i.e., brigimadlin or doxorubicin, will be produced for patients in the treated set. A descriptive summary of **total duration of cross-over treatment**, i.e., brigimadlin, will be produced for patients in the cross-over set. **Time to permanent discontinuation of trial medication** and **time to permanent discontinuation of cross-over treatment** will be analyzed using Kaplan-Meier methods.

- For patients who have *permanently discontinued* trial medication (or cross-over treatment):

  Duration of treatment [days] = date of last administration of trial medication (or cross-over treatment) – date of first administration of trial medication (or cross-over treatment) + 1.

- For patients who have *not yet permanently discontinued* trial medication (or cross-over treatment):

  Duration of treatment (censored) [days] = date of data cut-off – date of first administration of trial medication (or cross-over treatment) + 1.

Total duration of treatment with trial medication (or cross-over treatment) therefore also includes the time(s) when trial medication (or cross-over treatment) is temporarily delayed and subsequently reintroduced.

For patients treated with doxorubicin, the time period after first administration of subsequent brigimadlin is not included in the total duration of treatment with doxorubicin.

Cumulative dose of doxorubicin

A descriptive summary of the **cumulative dose of doxorubicin** (with units of mg/m^2) will be produced for patients treated with doxorubicin. Per protocol, the maximum cumulative dose of doxorubicin is 450 mg/m^2.

Dose reduction

The **frequency of patients with dose reduction of trial medication**, i.e., brigimadlin or doxorubicin, will be produced for patients in the treated set. The **frequency of patients with dose reduction of cross-over treatment**, i.e., brigimadlin, will be produced for patients in the cross-over set.

The **first and second dose reduction** will be analyzed descriptively for the treated set and cross-over set.

Time to first dose reduction will be analyzed descriptively and using Kaplan-Meier methods.

- For patients *with a dose reduction* of trial medication (or cross-over treatment):

  Time to first dose reduction [days] = date of first administration of the reduced dose of trial medication (or cross-over treatment) – date of first administration of trial medication (or cross-over treatment) + 1.

- For patients who *discontinue* trial medication (or cross-over treatment) *without a dose reduction*:

  Time to first dose reduction (censored) [days] = date of last administration of trial medication (or cross-over treatment) – date of first administration of trial medication (or cross-over treatment) + 1.

- For patients who have *not yet permanently discontinued* trial medication (or cross-over treatment) at the time of analysis and *without a dose reduction*:

  Time to first dose reduction (censored) [days] = date of data cut-off – date of first administration of trial medication (or cross-over treatment) + 1.

For patients treated with doxorubicin, the time period after first administration of subsequent brigimadlin is not included in the analyses of dose reduction of doxorubicin.

Dose delay

Two types of dose delay will be considered:

1) **Dose delay due to AE**

   All instances where a site has answered "Was there a delay in planned administration due to AE?" with "Yes" in the eCRF will be considered a case of dose delay due to AE.

2) **Dose delay >3 days due to any reason**

   Note: As per CTP, a window (+3 d) was acceptable for scheduling Day 1 of each cycle after Cycle 1.

   First, the time between all pairs of consecutive treatment cycles will be calculated, i.e., the time between Cycles 1 and 2, the time between Cycles 2 and 3, ..., the time between the last cycle and the data cut-off date. Any of these times that are >24 days will be considered cases of dose delay >3 days due to any reason, even if the site did not answer "Was there a delay in planned administration due to AE"? with "Yes" in the eCRF.

The following analyses will be performed for both types of dose delay.

The **frequency of patients with dose delay of trial medication**, i.e., brigimadlin or doxorubicin, will be produced for patients in the treated set. The **frequency of patients with dose delay of cross-over treatment**, i.e., brigimadlin, will be produced for patients in the cross-over set.

The **first cycle of dose delay** will be analyzed descriptively for the treated set and cross-over set.

Duration of longest dose delay will be analyzed using Kaplan-Meier methods. Patients without a dose delay will not be included in the analyses of duration of longest dose delay.

- For patients with a delayed dose of trial medication (or cross-over treatment) that *has* been administered:

  Duration of dose delay [days] = date of administration of the delayed dose of trial medication (or cross-over treatment) – planned date of administration of the delayed dose of trial medication (or cross-over treatment).

- For patients with a delayed dose of trial medication (or cross-over treatment) that *has not* been administered:

  Duration of dose delay (censored) [days] = date of data cut-off – planned date of administration of the delayed dose of trial medication (or cross-over treatment).

The planned date of a delayed dose is defined as the actual date of the latest administration of trial medication (or cross-over treatment) prior to dose delay + 21 days. The duration of longest dose delay is the largest calculated value regardless of censoring.

Relative dose intensity (RDI)

A descriptive summary of the relative dose intensity (RDI) of brigimadlin and doxorubicin, respectively, will be performed for the treated set. In general, RDI is the sum of actual doses divided by the sum of expected doses, i.e.,

$$\text{Relative dose intensity (\%)} = \frac{\text{Sum of actual doses}}{\text{Sum of expected doses}} \times 100\%.$$

Since brigimadlin is administered once every three weeks for as many cycles as deemed appropriate based on the investigator and the protocol, and doxorubicin has a maximum cumulative dose defined in the protocol, the calculation of RDI of brigimadlin differs from the calculation of RDI of doxorubicin.

The **RDI of brigimadlin** will be analyzed *up to the first six months* and *in the entire treatment period*. The numerator (i.e., the sum of actual doses of brigimadlin) is equal to the total sum of doses of brigimadlin actually received (either up to the first six months, or across the entire treatment period of brigimadlin, respectively) after the date of first administration of brigimadlin in mg. The denominator (i.e., the sum of expected doses of brigimadlin) is calculated as follows.

$$\text{Sum of expected doses}_{\,brigimadlin} = \max\left[\textbf{ceiling}\left(\frac{\text{Day}_{last\ trt}}{21\ \text{days}}\right), n_{cycles}\right] \times \text{Starting dose}$$

- For the **RDI of brigimadlin up to the first six months**:
  - $\text{Day}_{last\ trt}$ = analysis day of the last treatment date of brigimadlin *up to the first six months* relative to the date of first administration of brigimadlin,
  - n_{cycles} = total number of cycles of brigimadlin actually received *up to the first six months* after the date of first administration of brigimadlin.
- For the **RDI of brigimadlin in the entire treatment period**:
  - $\text{Day}_{last\ trt}$ = analysis day of the last treatment date of brigimadlin *across the entire treatment period of brigimadlin* relative to the date of first administration of brigimadlin,
  - n_{cycles} = total number of cycles of brigimadlin actually received *across the entire treatment period of brigimadlin* after the date of first administration of brigimadlin.
- For patients who increase their dose from brigimadlin 30 mg to 45 mg q3w, the increase in dose will be considered as an expected dose change that does not yield an RDI > 100%.

The **RDI of doxorubicin in the entire treatment period** will be analyzed as follows. The numerator (i.e., the sum of actual doses of doxorubicin) is equal to the total cumulative dose of doxorubicin actually received in mg/m^2. The denominator (i.e., the sum of expected doses of doxorubicin) is always equal to the maximum cumulative dose of 450 mg/m^2.

7.8 SAFETY ANALYSIS

All safety analyses described below will be performed on the treated set for the interim CTR and the final CTR. For the treated set, the displayed treatment groupings will be: “Brigimadlin 30 mg q3w”, “Brigimadlin 45 mg q3w”, “Brigimadlin Total”, and “Doxorubicin”. Some of the safety analyses described below may also be performed for the cross-over set for the interim CTR and the final CTR. For the cross-over set, the displayed treatment groupings will be: “Brigimadlin 30 mg q3w”, “Brigimadlin 45 mg q3w”, and “Brigimadlin Total”.

7.8.1 Adverse Events

AEs will be coded using the MedDRA coding dictionary. The most recent version that still allows the team enough time to generate and validate the displays will be used.

Severity of AEs will be scaled according to Common Terminology Criteria for Adverse Events (CTCAE) version 5.

The analyses of AEs will be descriptive in nature. Unless otherwise noted, analyses of AEs will be based on the number of patients with AEs and not on the number of AEs. Overlapping or adjacent AE records, with start dates at or after the first treatment administration, will be linked if they have the same MedDRA preferred terms and considered as describing the same AE episode.

The analysis of AEs will be based on the concept of treatment emergent AEs. That means that all AEs with an onset between first drug intake and last drug intake + REP (i.e., 30 days) will be assigned to the initial brigimadlin dose or doxorubicin. All AEs with an onset before first drug intake will be assigned to 'screening', and all AEs with an onset after the REP will be assigned to 'post-treatment' (for listings only). For details on the treatment definition, see Section 6.1.

An overall summary of AEs will be presented.

Frequency tables of patients with AEs summarized by treatment, primary SOC, PT, and maximum CTCAE grade will be produced. Frequency tables of patients with AEs summarized by treatment, grouped PT (see Table 7.8.1: 1) or PT, and maximum CTCAE grade will also be produced. SOCs, grouped PTs, PTs, and PTs within SOCs, will be sorted by descending frequency of AEs across all treatment groups being analyzed.

The following groupings of CTCAE grading will be used: Frequency tables of patients with AEs with maximum investigator reported CTCAE grading displayed as "All Grades", "Grade >= 3", "Grade 1", "Grade 2", "Grade 3", "Grade 4", "Grade 5".

AEs with a missing investigator reported CTCAE grade will be displayed under the category of "All Grades", and an explanatory footnote will be added. No "Missing Grade" category will be displayed.

Frequency tables of patients with AEs with an onset during the on-treatment period will be produced for:

- Patients with AEs
- Patients with investigator reported drug-related AEs
- Patients with SAEs
- Patients with investigator reported drug-related SAEs
- Patients with AEs of Special Interest (as defined in Section 5.2.6.1 of the CTP, not included in the interim CTR)
- Patients with AEs leading to dose reduction of trial medication
- Patients with AEs leading to dose delay of trial medication
- Patients with AEs leading to permanent discontinuation of trial medication

- Patients with AEs in user-defined AE categories (UDAECs) (with the latest version defined in Table 7.8.1: 1 below; UDAECs are defined on project level, and the latest version available in the project Safety SAP at the time of analysis will be used)
- Patients with AEs leading to death (no CTCAE grade required)

Table 7.8.1: 1 User-defined AE categories (UDAECs)

Name of user-defined category	MedDRA PTs to include in the category	Use as grouped PT in standard AE tables
Haematological AEs		
Thrombocytopenia+	- Include PTs under SMQ: 20000031 Haematopoietic thrombocytopenia (Narrow)	Yes
Anaemia+	- Include PTs under SMQ: 20000029 Haematopoietic erythropenia (Narrow) - Anaemia - Haemoglobin decreased - Leukoerythroblastic anaemia - Normochromic anaemia - Normochromic normocytic anaemia - Normocytic anaemia	Yes
Neutropenia+	- Febrile neutropenia - Neutropenia - Neutropenic infection - Neutropenic sepsis - Neutrophil count decreased - Radiation neutropenia - Band neutrophil count decreased - Band neutrophil percentage decreased - Cyclic neutropenia - Idiopathic neutropenia	Yes
Lymphopenia+	- Lymphocyte count decreased - Lymphopenia - B-lymphocyte count decreased - T-lymphocyte count decreased - CD19 lymphocytes decreased	Yes
Other Leukopenia+	- Include PTs under SMQ: 20000030 Haematopoietic leukopenia (Narrow) - Exclude PTs already in Neutropenia+ and Lymphopenia+	Yes
Other Cytopenias+	- Include PTs under SMQ 20000028 Haematopoietic cytopenias affecting more than one type of blood cell (Narrow)	Yes
Gastrointestinal AEs		

Nausea	- Nausea	No
Vomiting+	- Vomiting - Vomiting projectile - Retching	Yes
Diarrhoea	- Diarrhoea	No
Other Gastrointestinal AEs+	- Include PTs under SMQ: 20000137 Gastrointestinal nonspecific inflammation and dysfunctional conditions (Narrow) - Exclude PTs already in Nausea, Vomiting+, and Diarrhoea - Exclude PT "Non-cardiac chest pain"	No
Haemorrhage		
Haemorrhage+	- Blood urine - Blood urine present - Gastric occult blood positive - Occult blood positive - Urinary occult blood - Urinary occult blood positive - Include PTs under SMQ: 20000039 Haemorrhage terms (excl laboratory terms) (Narrow)	No
Infections		
Infections+	- Include PTs under SOC: Infections and infestations - SMQ 20000235 Opportunistic infections (Narrow) - SMQ 20000234 Sepsis (Narrow) - SMQ 20000231 Infective pneumonia (Narrow)	No
Fatigue		
Fatigue+	- Fatigue - Cancer fatigue - Asthenia - Lethargy - Malaise	Yes

+ indicates grouped PTs.

UDAECs are defined on project level, and the latest version available in the project Safety SAP at the time of analysis will be used.

Listings of AEs will be displayed by patient. The actual dose of brigimadlin or doxorubicin administered at the day of AE onset will be derived and included in the listings. AEs will be reported with start day and end day as calculated from the first day of treatment with brigimadlin or doxorubicin. For listings displaying AEs during screening and after the on-treatment period, the start and end days are calculated from the start of the respective analysis period.

In addition, hematological and gastrointestinal AEs will be analyzed over time. Frequency of patients will be summarized by the maximum CTCAE grade and treatment cycle, and by AE

outcome at the end of the cycle. Time to first onset will also be summarized, with Kaplan-Meier estimates provided. Time to first onset for a patient with no observed events will be censored at the earliest of trial medication stop date + 30 days, death date, or data cut-off date. Therapies for the AEs will be summarized. Table 7.8.1: 2 defines the categories of therapies of particular interest. (Categories of therapies of particular interest are defined on project level, and the latest version available in the project Safety SAP at the time of analysis will be used.)

Table 7.8.1: 2 Therapies of particular interest for selected AE grouped terms

Therapy category	Definition by ATC code or MedDRA PT	Selected AE grouped terms
Antiemetics	- A04 Antiemetics and antinauseants	Nausea, Vomiting+
Antidiarrheals	- A07D Antipropulsives	Diarrhoea
GCSF	- L03AA Colony stimulating factors	Neutropenia+
Red blood cell transfusion	- B05AX01 Erythrocytes - MedDRA PT: Red blood cell transfusion	Anaemia+
Other antianemic therapies	- B03 Antianemic preparations	Anaemia+
Platelet transfusion	- B05AX02 Thrombocytes - MedDRA PT: Platelet transfusion	Thrombocytopenia+
Interleukins	- L03AC Interleukins	Thrombocytopenia+, Neutropenia+
Thrombopoietin receptor agonist	- B02BX Other systemic hemostatics	Thrombocytopenia+
Blood transfusion	- B05AX03 Blood plasma - MedDRA PT: Transfusion	Anaemia+, Thrombocytopenia+, Neutropenia+
Antiinfectives for systemic use	- J01 Antibacterials for systemic use - J02 Antimycotics for systemic use - J04 Antimycobacterials - J05 Antivirals for systemic use - J06 Immune sera and immunoglobulins - J07 Vaccines	Infections+

Categories of therapies of particular interest are defined on project level, and the latest version available in the project Safety SAP at the time of analysis will be used.

[redacted]

Exposure-adjusted analyses of adverse events

AE analyses will be presented in the same table unadjusted and also exposure-adjusted to calculate the incidence rate.

The incidence rate (per 100 patient-years) of a selected AE (also known as the incidence density rate or person-time incidence rate) is defined as the number of patients experiencing the AE per treatment group during time at risk divided by the total time of patients at risk in that treatment group to contribute an event to the analysis multiplied by 100 (per 100 patient-years), where

Time at risk [patient-years] = (date of onset of AE – study drug start date + 1) / 365.25

If no treatment emergent AE occurred for a patient, then the time at risk will be censored at the earliest of trial medication stop date + 30 days, death date, or data cut-off date.

Incidence rate [1/100 Patient-years (pt-yrs)] = 100 * number of patients with AE / Total AE-specific time at risk [years]

7.8.2 Laboratory data

The analyses of laboratory data will be descriptive in nature and based on BI standards (5). Derivations of CTCAE grade for laboratory values will follow BI standards (6).

Only patients with baseline and at least one on-treatment value available will be included in the analysis of an individual laboratory parameter.

Laboratory measurements taken up to 30 days after the last administration of trial medication will be considered as on-treatment.

CTCAE grades for applicable laboratory parameters will be calculated according to CTCAE version 5. Table 7.8.2: 1, Table 7.8.2: 2, and Table 7.8.2: 3 list the labs collected in the trial along with categorizations based on importance and planned analyses. Rules used to determine possible clinical significance abnormalities are also listed in the tables.

For primary laboratory tests listed in Table 7.8.2: 1, the following analyses will be performed:

- Descriptive statistics, including change from baseline, will be provided for the following values:

  - Normalized laboratory value at baseline
  - Worst laboratory value on treatment
  - Last laboratory value on treatment, defined as the laboratory value of the last visit during the on-treatment period.

- For labs with CTCAE grading, shift tables of transitions of CTCAE grade from baseline to last value on treatment, from baseline to worst value on treatment, and from worst value to last value on treatment will be presented. For labs with no CTCAE grading, shift tables of transitions relative to reference ranges will be presented.
- With CTCAE version 5, the grading of some laboratory parameters can no longer be assigned a grade at baseline. Therefore, the CTCAE shift tables for ALT, AST, bilirubin, and creatinine will be replaced by shift tables of transitions relative to multiples of Upper Limit of Normal (ULN) values with the following categories:
  - ALT and AST: ≤ULN, >1 – 3×ULN, >3 – 5×ULN, >5 – 20×ULN, >20×ULN
  - Bilirubin: ≤ULN, >1 – 1.5×ULN, >1.5 – 3×ULN, >3 – 10×ULN, >10×ULN
  - Creatinine: ≤ULN, >1 – 1.5×ULN, >1.5 – 3×ULN, >3 – 6×ULN, >6×ULN
- Frequency of patients with possibly clinically significant abnormalities will also be presented.

For secondary laboratory tests listed in Table 7.8.2: 2, analyses will be limited to frequency of patients with possibly clinically significant abnormalities.

Table 7.8.2: 1 Primary laboratory tests

Lab Test Code	Lab Test Name	Direction of Interest	Potential Clinical Significance (CS) Rule†
ALT	Alanine Aminotransferase	High‡	CTCAE grade 2 or greater
AST	Aspartate Aminotransferase	High‡	CTCAE grade 2 or greater
BILI	Total Bilirubin	High‡	CTCAE grade 2 or greater
CREAT	Creatinine	High‡	> 1.5 × ULN and > baseline
HGB	Haemoglobin	Low	CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline
LYM	Lymphocytes	Low	CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline

NEUT	Neutrophils	Low	CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline
PLAT	Platelets	Low	CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline
RBC	Erythrocytes	Low*	< 3 × 10^12/L and < baseline
WBC	Leukocytes	Low	CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline

† If baseline value is missing, any comparison with baseline will not be required to meet the CS rule.
‡ Transition table will be based on multiples of ULN.
* No transition tables will be produced for RBC.

Table 7.8.2: 2 Secondary laboratory tests

Lab Test Code	**Lab Test Name**	**Direction of Interest**	**Potential Clinical Significance (CS) Rule†**
ALB	Albumin	Low	CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline
ALP	Alkaline Phosphatase	High	CTCAE grade 2 or greater
APTT	Activated Partial Thromboplastin Time	High	CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline
BASO	Basophils	High	> 0.2 × 10^9 L
BICARB	Bicarbonate	High‡, Low‡	High: > 32.0 mmol/L and > baseline Low: < 18.0 mmol/L and < baseline
CA*	Calcium	High, Low	CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline
CK	Creatine Kinase	High	CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline
CL	Chloride	High, Low	High: > 120.0 mmol/L and > baseline Low: < 80.0 mmol/L and < baseline
EOS	Eosinophils	High‡	> 1 × 10^9/L and > baseline
GLUC	Glucose	High‡, Low	High: > 10.0 mmol/L and > baseline Low: CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline

INR	Prothrombin Intl. Normalized Ratio	High‡	> 1.5 and > baseline
K	Potassium	High, Low‡	High: CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline Low: < 3 mmol/L and < baseline
LDH	Lactate Dehydrogenase	High	> 3 × ULN and > baseline
MCV	Ery. Mean Corpuscular Volume	High, Low	High: > 110 fL Low: < 65 fL
MONO	Monocytes	High, Low	High: > 0.92 x 10^9 L Low: < 0.12 x 10^9 L
PHOS	Phosphate	High, Low	High: > 1.7 mmol/L and > baseline Low: < 0.7 mmol/L and < baseline
PROT	Protein	Low	< LLN and < baseline
PT	Prothrombin Time (seconds)	High	> ULN and > baseline
PT	Prothrombin Time (%, activity measurement)	Low	< LLN and < baseline
SODIUM	Sodium	High, Low‡	High: CTCAE grade 2 or greater, with an increase of at least one CTCAE grade from baseline Low: < 130 mmol/L and < baseline
URATE	Urate	High‡	Female: > 600 umol/L and > baseline Male: > 650 umol/L and > baseline
UREA	Urea	High	> 1.5 × ULN and > baseline
UREAN	Urea Nitrogen	High	> 10 mmol/L and > baseline

† If baseline value is missing, any comparison with baseline will not be required to meet the CS rule.
‡CTCAE grades can only be differentiated by taking physiological consequences into account. Analysis of these laboratory parameters for the corresponding directions of interest will not be based on CTCAE grade.
*CTCAE grading based on corrected calcium.

Table 7.8.2: 3 Laboratory tests that will not be analyzed for the interim CTR

Lab Test Code	Lab Test Name
AMYLASEP	Amylase, Pancreatic
APTTSTND	Activated PTT/Standard
BILIDIR	Direct Bilirubin
BILIIND	Indirect Bilirubin
CKMB	Creatine Kinase MB

HCG	Choriogonadotropin Beta
HCT	Hematocrit
LIPASET	Lipase
MGB	Myoglobin
TROPON I	Troponin I
TROPON T	Troponin T

Urine analysis (including urine sediment), infectious serology testing, and tumor markers will also not be analyzed. For differential WBC counts, only absolute values will be analyzed; percentages will not be analyzed. For CKMB, neither mass concentration or activity measurement will be analyzed.

Differential white blood cell (WBC) count

Differential WBC counts (lymphocytes, monocytes, basophils, eosinophils, and neutrophils) are preferably measured in absolute values according to the protocol. In case a patient's differential blood count is reported as a percentage, the following conversion will be applied:

Differential [10^9/L] = WBC count [10^9/L] × Differential [%] / 100.

No conversion will be applied to reference ranges; BI standard reference ranges will be used.

Corrected calcium

Grading of hypocalcemia is based on corrected calcium in CTCAE version 5.0, as calcium can be falsely low if hypoalbuminemia is present. The following corrective calculation will be performed:

Corrected calcium [mmol/L] = Total calcium [mmol/L] + 0.02×(40g/L – Albumin [g/L]).

The reported reference range of total calcium will be used for analyses without corrections.

Liver function tests and potential Hy's law cases

The assessment of hepatic enzyme elevations will include frequencies of patients falling into the following categories:

- Maximum ALT (AST): >1 – 3×ULN, >3 – 5×ULN, >5 – 20×ULN, >20×ULN
- Maximum total bilirubin: >1 – 1.5×ULN, >1.5 – 3×ULN, >3 – 10×ULN, >10×ULN
- Maximum ALP: >1 – 2.5×ULN, >2.5 – 5×ULN, >5 – 20×ULN, >20×ULN
- Normal ALT (AST) at baseline and: ALT (AST) ≥3×ULN with total bilirubin ≥2×ULN, ALT (AST) ≥10×ULN
- ALT (AST) >ULN – 3×ULN at baseline and: ALT (AST) ≥3×baseline with elevated total bilirubin, ALT (AST) ≥5×baseline
- ALT (AST) >3×ULN – 5×ULN at baseline and: ALT (AST) ≥2×baseline with elevated total bilirubin, ALT (AST) ≥3×baseline

Where applicable for defining the above, events can occur in any order but must occur within 30 days of each other with an increase of either ALT or AST being the trigger for evaluations. Along with summaries of the above, eDISH plots of maximum ALT (AST) versus total bilirubin (maximum within 30 days after maximum ALT (AST) elevation) will be produced.

Laboratory values of special interest

In depth analyses will be performed for laboratory evaluations associated with the special safety topic on haematology. Analysis will be based on normalized values when pooling data from local laboratories.

Descriptive statistics of nadir by actual cycle will be presented for the following laboratory values: platelets, leukocytes, neutrophils, and hemoglobin.

For absolute neutrophil counts (ANC) and platelet counts, first treatment cycle with a value below a threshold will be provided. The thresholds of interest are 1.5×10^9/L, 1.0×10^9/L, and 0.5×10^9/L for neutrophil counts, and 100×10^9/L, 50×10^9/L, and 25×10^9/L for platelet counts. In addition, the following time-to-event values will be summarized:

- Time to first occurrence of a lab value below threshold, in terms of study day with Day 1 being the day of first administration of trial medication. Time will be censored on the day of the last value on treatment for patients without an event.
- Time from first occurrence of a lab value below threshold to first subsequent recovery above protocol-specified thresholds, with Day 1 being the day of the first drop below threshold. Time will only be derived for patients with a drop below threshold, and will be censored on the day of the last value on study for patients with no observed recovery to protocol-specified $\geq100\times10^9$/L for platelet counts or $\geq1.5\times10^9$/L for ANC.
- Time to overall nadir among patient with some values below threshold, in terms of study day with Day 1 being the day of first administration of trial medication. The analysis will include only on treatment, and conduct for the highest threshold only, i.e. 1.5×10^9/L for neutrophil counts, and 100×10^9/L for platelet counts.
- Time from overall nadir to first subsequent recovery above threshold among patients with some values below the threshold. Time will be censored on the day of last value on study for patients with no observed recovery to protocol-specified $\geq100\times10^9$/L for platelet counts or $\geq1.5\times10^9$/L for ANC.

Kaplan-Meier curves and estimates will be provided. For time to first occurrence and time to first subsequent recovery, crude summary statistics of the time-to-event among patients with an event and recovery respectively will also be provided.

[redacted]

7.8.3 Vital signs

Only descriptive statistics are planned for this section of the interim CTR and the final CTR.

7.8.4 ECG

Electrocardiogram (ECG) results will only be listed in the interim CTR and the final CTR; any clinically significant abnormalities will be recorded as either a concomitant diagnosis or an AE.

7.8.5 Others

Analysis for dose selection

The interim analysis for dose selection will be conducted during Phase II. More details are in Section 7.2.8 of the CTP, the DMC SAP, and a separate PK/PD SAP.

In addition, an updated dose selection report will be produced at the time point of the database snapshot for the primary analysis of PFS (based on blinded central independent review). In this report, safety and efficacy analyses supporting dose selection will be based on data collected from the two brigimadlin arms in the Phase II part of trial 1403-0008, where patients were concurrently randomized to 30 mg q3w and 45 mg q3w brigimadlin treatment arms.

8. TIME POINT OF RELEASE OF TREATMENT INFORMATION

Handling of individual patient treatment information

This is an open-label trial. Therefore, no blinding will apply for patients, investigators, and personnel at the sites involved in trial conduct. The access to the (planned) randomization code list will be kept restricted until it is released for analysis. The treatment information for an individual patient will be available in the trial database after their randomization.

Handling of aggregated treatment information

While the trial is in progress and prior to submission of the "Data Ready to be Unblinded and/or Final Trial Closure Notification" (RUN) form (i.e., prior to the data being declared ready for unblinded aggregate analysis by the trial team), access to tabular results of trial outcomes by treatment will not be made available to patients, investigators, personnel at the sites involved in trial conduct, the trial statistician, clinical team, or members of the steering committee (unless the DMC advises otherwise). Any tabular displays of trial outcomes generated by the trial team prior to submission of the RUN form will be produced using mock (dummy) treatment codes. The DMC (which includes the DMC statistician) and the independent statistician (iSTAT) will be unblinded.

At the time point of the database snapshot for the primary analysis of PFS (based on blinded central independent review), the actual treatment codes will be used to unblind the data for analysis by the trial team. Details about the data lock point, data access, and logistics for the interim analysis for dose selection in the Phase II part are described in a separate "Interim

Analysis for Dose Selection Logistics and Access Plan". Details about the data lock point, data access, and logistics for the interim futility analysis are described in Section 10.1 below.

"Fast-track" database snapshot for PK/PD analyses for interim CTR

Prior to the time point of interim database snapshot for the primary analysis of the primary endpoint PFS (based on blinded central independent review) described below, an earlier "fast-track" database snapshot is planned for PK/PD analyses described in a separate PK/PD analysis plan. These initial analyses will be performed by authorized TMCP analysts in collaboration with the external vendor Pharmetheus solely to prepare and set up the analysis programs in advance. Access to unblinded data, programs, and any interim results will be restricted to the authorized individuals specified in the Pharmetheus unblinding charter. The TSTAT will have already requested that the trial medication and randomization codes are released to the unblinded TMCP analysts by submitting the Treatment Information Release (TIR) form to the global Randomization Specialist (gRS) for the interim analysis for dose selection. The "fast-track" process is also mentioned in Section 7.2.7 of the CTP.

Time point of interim database snapshot for the primary analysis of the primary endpoint PFS (based on blinded central independent review) for the interim CTR

The primary analysis of PFS (based on blinded central independent review) will occur during Phase III and is estimated to take place approximately 23 months after the start of enrollment of the Phase II part of the trial.

Once the required number of PFS events (as specified in Section 7.4.1) have occurred and all corresponding data have been entered and cleaned to the level documented in the "Data Delivery Request" (DDR) form, the data will be declared ready for aggregate analysis via the RUN form. Then, the planned treatment information (i.e., randomization codes) will be released into the analysis database.

Time point of final database lock for the primary analysis of the secondary endpoint OS for the final CTR

The primary analysis of OS will occur at the end of Phase III. After the interim database snapshot for the primary analysis of the primary endpoint PFS (based on blinded central independent review), data collection for the primary OS analysis will continue into the trial database. Once the required number of OS events (as specified in Section 7.5.2) have occurred, all trial data collection has been completed, and all data have been entered and cleaned as documented on the RUN form, a final database lock for the final CTR will be performed.

9. REFERENCES

1.	*CPMP/ICH/363/96:* "Statistical Principles for Clinical Trials", ICH Guideline Topic E9, Note For Guidance on Statistical Principles for Clinical Trials, current version.
2.	*BI-VQD-12161_ 30-476:* "TMCP Data Analysis".
3.	*BI-KMED-TMCP-HTG-0025:* "Standards and processes for analyses performed within Clinical Pharmacokinetics/Pharmacodynamics", current version.
4.	*BI-KMED-TMCP-MAN-0014:* "Noncompartmental PK/PD Analyses of Clinical Studies", current version.
5.	*BI-KMED-BDS-HTG-0042:* "Handling, Display and Analysis of Laboratory data", current version.
6.	*BI-KMED-BDS-HTG-0036*: "CTCAE Grading for Laboratory Values", current version.
R21-3910	Di Scala L, and Glimm E. "Time-to-event analysis with treatment arm selection at interim." Statistics in medicine 30.26 (2011): 3067-3081
R14-1929	Cocks K, King MT, Velikova G, de Castro G Jr, Martyn St-James M, Fayers PM, Brown JM. Evidence-based guidelines for interpreting change scores for the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. Eur J Cancer. 2012 Jul;48(11):1713-21
R22-3804	Basch E, LJ Rogak, and AC Dueck, Methods for implementing and reporting patient-reported outcome (PRO) measures of symptomatic adverse events in cancer clinical trials. 2016. 38(4): 821-830
R09-6372	Brookmeyer R, Crowley J. A confidence interval for the median survival time. Biometrics 1989; 38; 29-41
R14-1197	Lehmacher W, Wassmer G. Adaptive sample size calculations in group sequential trials. Biometrics 1999; 55; 1286-1290
R14-2334	Bretz F, Koenig F, Brannath W, Glimm E, Posch M. Adaptive designs for confirmatory clinical trials. Stat Med 2009; 28(8), 1181-1217
R15-0928	Wassmer G. Planning and analyzing adaptive group sequential survival trials. Biometr J 2006; 48(4); 714-729
R15-5218	Jenkins M, Stone A, Jennison C. An adaptive seamless phase II/III design for oncology trials with subpopulation selection using correlated survival endpoints. Pharm Stat 2011; 10(4); 347-356

R15-5971	Robins JM, and Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Commun. Stat Theory Methods; 1991; 20(8); 2609-2631
R07-2064	Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Published by: European Organisation for Research and Treatment of Cancer, Brussels 2001
R09-1299	Greenland S, and Robins JM. (1985). Estimation of a Common Effect Parameter from Sparse Follow-Up Data. Biometrics 41:55-68
R23-4188	Uschner D, Sverdlov O, Carter K, Chipman J, Kuznetsova O, Renteria J, Lane A, Barker C, Geller N, Proschan M, Posch M, Tarima S, Bretz F, Rosenberger WF. (18 Oct 2023): Using Randomization Tests to Address Disruptions in Clinical Trials: A Report from the NISS Ingram Olkin Forum Series on Unplanned Clinical Trial Disruptions. Statistics in Biopharmaceutical Research
R23-4187	Grassano L, Ranzato G, Pellegrini M, Costantini M. (2023). Re-randomization tests as sensitivity analyses to confirm immunological noninferiority of an investigational vaccine: Case study. Pharmaceutical Statistics. 22(3):570-576